Platform processes for vaccine production:

Development of a universal influenza vaccine using *Pichia pastoris*

A thesis submitted to University College London for the degree of

DOCTORATE OF PHILOSOPHY

By

Benjamin A. F. Bláha

The Advanced Centre for Biochemical Engineering, Department of Biochemical Engineering, UCL,

Bernard Katz Building, Gordon Street, WC1E 7JE, United Kingdom

Declaration

I, Benjamin André František Bláha, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Influenza virus remains a familiar threat to public health and continues to challenge the scientific community. Most challenging is responding to the evolutionary adaptability of influenza virus, which often hinders effective prevention or treatment. Furthermore, current production methods are logistically and technically inadequate to cope with pandemic surges, leaving a considerable number of individuals unprotected.

Tandem Core Virus-Like Particles (VLPs), expressed in *Pichia pastoris*, offer an exciting proposition to create a platform process for a universal influenza vaccine. However, as with all novel concepts, characterisation of this technology was required and process methodologies were developed accordingly.

Initially**,** critical process parameters, associated with production of novel VLP constructs, were identified. In doing so, a need for a robust, high-throughput miniaturised fermentation platform was recognised. To accommodate such an approach, a high-throughput, non-contact, automated, small-scale, scalable disruption tool was developed to extract VLPs from *P. pastoris*. Development and optimisation of this method led to matching and even outperforming High Pressure Homogenisation performance.

Having developed the prerequisite tools for miniaturised upstream process development, the use of microtitre plates for studying heterologous protein expression was investigated. This proved to be challenging, predominantly due to

reduced mass transfer capacity imposed by microtitre plates and the requirement of controlled methanol feeding. The use of pectin digest as an alternative indirect induction agent was studied. It was found that the use of this novel media resulted in the expression of heterologous protein, however, this effect was most likely a result of methanol liberation during digest preparation. Due to the requirements of a tightly controlled induction process, it was found that, microtitre plates provided an unsuitable platform for rapid upstream process development for methylotrophic yeast.

As an alternative to microtitre plates, the effects of variance of previously identified critical process parameters were studied to a greater extent using miniaturised bioreactors, initially with a simplified, non-epitope-exposing, variant of Tandem Core VLP. Metabolic responses, such as final biomass concentration, required significantly different optimum operating conditions than product expression responses. It was also found that maximising expression of heterologous protein was not synonymous to maximising the production of VLPs.

After characterising VLP expression in miniaturised bioreactors, these findings were translated to industrially relevant universal influenza vaccine production scenarios. This was accomplished by investigating the effects of scale-up and variation of epitope inserts that could be utilised as a universal influenza antigens. Following this, the scale-up of fermentation processes and was studied. It was found that mixed induction with glycerol feeding not only provided a better basis for scale-up but also resulted in higher product titres for more complex VLP constructs.

Impact statement

Characterising the expression of Tandem Core Virus-Like Particles in *P. pastoris* has gained:

- Increased understanding in the potential and limitations of Tandem Core technology as an epitope scaffold for customisable vaccine development. This knowledge will pave the way, not only to a universal influenza vaccine production process, but to a universal production process for vaccines in general.
- Optimised and scalable processes for Tandem Core VLP production, allowing for scalable and maximum vaccine production capacity. This is particularly beneficial in the event of a pandemic surge.
- Increased understanding of metabolic effects in *P. pastoris* fermentations. These findings could translate well to process development of many biopharmaceuticals
- Multiple response models and correlations that can be used to predict performance of cell disruption using high pressure homogenisers. This knowledge extends to extraction of VLPs and biopharmaceuticals other than VLPs.
- Development of a high-performance, high-throughput, non-contact, automated, small-scale, scalable disruption tool for microbial bioprocess development. The utility of this work means that, for microbial cultures with sizeable cell walls, such as *P. pastoris*, an efficient scale-down tools is now available adding to the processing tools set. This can be used for a wide range of applications including buffer development, strain selection, fermentation process development and whole bioprocess integration.

“We are not going in circles, we are going upwards.

The path is a spiral; we have already climbed many steps.”

-Hermann Hesse, Siddhartha

Acknowledgments

Pursuing a doctorate degree has been one of the richest learning experiences of my life. I feel incredibly fortunate for having had this opportunity.

I am greatly indebted to my supervisor, Dr Tarit Mukhopadhyay, for his constant support, guidance and constructive criticism. My deepest gratitude and appreciation are expressed to him.

I would like to thank my parents for their constant encouragement and for a lifetime of invaluable support, enabling me to pursue a career is science and engineering.

I am grateful to my partner, Rhian, whose daily support made heavy workloads considerably lighter.

I would also like to thank Dr Stephen Morris and Dr Michael Sulu for their help throughout my time in the laboratory.

Finally, I am grateful for this experience and for the friendships built along the way.

It’s been a blast.

Chapter 1 Introduction 32

1.1 Background 32

1.2 Influenza Virology 33

1.2.1 Viral composition 33

1.3 Mechanism of infection 34

1.4 Viral taxonomy 37

1.4.1 Influenza genera 38

1.4.2 Species: antigenic drift and antigenic shift 38

1.4.3 Nomenclature 40

1.4.4 Seasonal influenza 40

1.4.5 Pandemic influenza 41

1.5 Immunology 42

1.5.1 The innate immune response 43

1.5.2 The adaptive immune response 43

1.5.3 Immune response to influenza 45

1.5.4 Immunogenicity 46

1.6 Pharmaceutical interventions against influenza 47

1.6.1 Antiviral drugs 47

1.6.2 Vaccines 48

1.6.3 Adjuvants 51

1.7 Vaccine production 51

1.7.1 Fertilised chicken egg-derived vaccine production 51

1.7.2 Vaccine production in mammalian cells 54

1.7.3 Vaccine production in *Escherichia coli* 56

1.7.4 Vaccine production in insect cells 59

1.8 Tandem Core technology 62

***1.9 Pichia pastoris*..... 65**

1.9.1 Bioprocessing – fermentation 66

1.10 Thesis aims 68

Chapter 2 Materials and methods 72

2.1 Chemicals and media 72

2.2 Yeast strains 72

2.2.1 Cell bank preparation 73

2.3 Fermentations 75

2.3.1 Microwell fermentations 75

2.3.2 Shake flask cultures 77

2.3.3 Mut^S 250 millilitre bioreactor fermentations 77

2.3.4 Mut^S 7 litre bioreactor fermentations 81

2.3.5 Mut^S 30 litre bioreactor fermentations 81

2.3.6 Mut^+ 7 litre bioreactor fermentations 83

2.4 Preparative methods 84

2.4.1 Small-scale harvesting 84

2.4.2 Large-scale harvesting 85

2.4.3 Cell disruption with High Pressure Homogenisation (HPH) 85

2.4.4 Cell disruption with sonication 86

2.4.5 Cell disruption with Adaptive Focused Acoustics (AFA) 86

2.4.6 Cell disruption with enzymatic lysis 87

2.4.7 Chemical lysis 87

2.4.8 Cell disruption with YeastBuster™ 87

2.4.9 Bead lysis 88

2.4.10 Clarification 88

2.4.11 Preparative Size Exclusion Chromatography (SEC) 89

2.5 Analytical methods 90

2.5.1 Biomass quantification using Optical Density (OD) 90

2.5.2 Biomass quantification using Wet Cell Weight (WCW) 90

2.5.3 Biomass quantification using Dry Cell Weight (DCW) 91

2.5.4 Total protein analysis with Bicinchoninic Acid (BCA) assay 91

2.5.5 Total protein analysis with Nanodrop 1000 92

2.5.6 Nucleic acid analysis with Nanodrop 1000 92

2.5.7 Product quantification using Green Fluorescent Protein (GFP) 92

2.5.8 Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE) ... 93

2.5.9 Western blotting 93

2.5.10 Dot blotting 94

2.5.11 Immunostaining 95

2.5.12 Particle Size Distribution analysis with static light scattering 95

2.5.13 Methanol analysis using High Performance Liquid Chromatography (HPLC) 96

2.5.14 Neutral sugar analysis using High Performance Liquid Chromatography (HPLC) 96

2.5.15 Uronic acid analysis using High Performance Liquid Chromatography (HPLC) 97

2.5.16 Transmission Electron Microscopy (TEM) 97

2.5.17 Phase-contrast and fluorescence microscopy 97

2.5.18 Statistical analysis and Design of Experiments (DoE) 97

Chapter 3 Initial process characterisation of novel Tandem Core VLP constructs 99

3.1 Introduction 99

3.1.1 Hypotheses and outline 102

3.2 *Pichia pastoris Mut*$^{+}$ expressing Tandem Core protein with GFP insert 103

3.2.1 Bench-top fermentation of *Pichia pastoris* Mut^{+} 104

3.2.2 Investigating alternative modes of induction and primary recovery 108

3.3 *Pichia pastoris* Mut^{S} expressing Tandem Core protein with lysine inserts 114

3.3.1 Fermentation of *Pichia pastoris Mut*^{S} 116

3.3.2 Initial DSP of VLP material 117

3.3.3 Investigating process robustness 120

3.4 Summary 123

Chapter 4 Characterisation, miniaturisation and optimisation of primary recovery methods 126

4.1 Introduction 126

4.2 Methods of cell disruption 127

4.2.1 High Pressure Homogenisation 128

4.2.2 Miniaturised cell disruption platforms 129

4.2.3 Aim and objectives 130

4.3 Results 131

4.3.1 Buffer Studies 131

4.3.2 Investigating the effect of detergent type and quantity 133

4.3.3 Buffer characterisation using a DoE methodology 135

4.4 Homogenisation characterisation 140

4.5 Development of small-scale cell disruption methods 144

4.5.1 Enzymatic lysis with lyticase 144

4.5.2 Adaptive Focused Acoustics 147

4.6 Adaptive Focused Acoustics characterisation 150

4.6.1 Adaptive Focused Acoustics Screening 151

4.6.2 Adaptive Focused Acoustics Optimisation 152

4.7 Causes and effects of micronised debris 158

4.8 Process scaling and cell disruption mimicry 161

4.9 Summary 163

Chapter 5 Microscale upstream process platform development 165

5.1 Introduction 165

5.2 Aim and objectives 165

5.2.1 Batch-mode growth kinetics screening in microtitre plates 166

5.3 Assessing methanol toxicity 172

5.4 Induction screening in microtitre plates 174

5.4.1 Initial induction screening results 174

5.4.2 Second induction screening 178

5.5 Investigating alternative feeding with pectin and pectin derivatives 186

5.5.1 Micro fermentations of pectin digest 188

5.5.2 Micro fermentations of pectin monomers 190

5.6 Summary 194

Chapter 6 Characterisation of Tandem Core Virus-Like particle production . 196

6.1 Introduction 196

6.2 Aim and objectives 197

6.3 Experimental Approach 197

6.3.1 Metabolic effects studied in the miniature bioreactor system 198

6.3.2 Specific expression 199

6.3.3 Design of Experiments 201

6.3.4 Methodology 204

6.4 Metabolic responses 206

6.4.1 Effect of methanol uptake 206

6.4.2 Biomass 208

6.4.3 The impact of variables on the CER 211

6.4.4 Respiratory quotient 213

6.5 Specific expression of Tandem Core protein 216

6.5.1 Specific total core protein 216

6.5.2 Specific soluble core protein 218

6.5.3 Time course profiles of specific expression responses 219

6.5.4 Specific VLP-core protein material analysis 222

6.6 Response-specific optima 224

6.7 Summary 227

6.8 Discussion 228

Chapter 7 Scale-up and application of Tandem Core technology to universal influenza vaccine candidates 230

7.1 Introduction 230

7.2 Aims and objectives 230

7.3 Variation of inserts with universal influenza epitopes 231

7.4 Effects of scale-up 239

7.4.1 Challenging scale-up of pure methanol cultures 240

7.4.2 Investigating scalability of mixed feeding induction 243

7.5 Impact of varying inserts on large-scale fermentation 246

7.6 Summary 249

Chapter 8 Conclusions and future work 251

8.1 Conclusions 251

8.2 Future work 253

Chapter 9 References 259

Chapter 10 Appendix 291

10.1 Supplementary data (Chapter 2) 291

10.1.1 Media components 291

10.1.2 Cell bank data 292

10.2 Supplementary data (Chapter 3) 294

10.2.1 Fermentation graphs 294

10.3 Supplementary data (Chapter 4) 297

10.3.1 Chemical lysis with YeastBuster™ 297

10.3.2 Sonication 299

10.4 Supplementary data (Chapter 5) 301

10.4.1 Theory of the effect of working volume on growth and expression 301

10.4.2 Design space for growth kinetics screening in microtitre plates 303

10.4.3 Calculation for correlation between incubation time and working volume 303

10.4.4 Design space and model summaries for induction screenings in microtitre plates 304

10.4.5 Reaction mechanism of d-galacturonic acid and methanol 308

10.5 Supplementary data (Chapter 6) 309

10.5.1 Volumetric total core protein 309

10.5.2 Volumetric expression responses..... 310

10.6 Supplementary data (Chapter 7) 314

List of figures

Figure 1.1 Schematic of influenza virus 34

Figure 1.2 Influenza virus lifecycle 35

Figure 1.3 Antigenic shift in a swine host 42

Figure 1.4 Downstream process of egg and mammalian cell derived vaccines 53

Figure 1.5 Overview of the production process for the manufacturing of M2e fusion proteins. 58

Figure 1.6 Overview of the production process for the manufacturing of rHA. 60

Figure 1.7 Tandem Core Technology overview 64

Figure 1.8: Typical fermentation of *P. pastoris*. 67

Figure 2.1 Design of the tHBc-k1,k1 construct. 73

Figure 2.2 Overview of the 250mL ambr® modular bioreactor system 80

Figure 2.3 Typical size exclusion profile for tHBc-k1, k1 90

Figure 3.1 Fermentation data of *P. pastoris* Mut^{+} expressing tHBc-GFP,e 106

Figure 3.2 Western blots of reduced gels of collected and processed fermentation samples 107

Figure 3.3 Specific tHBc-GFP,e expression of harvested material from three fermentations, each with different modes of induction 110

Figure 3.4 Fermentation sample analysis 110

Figure 3.5 Time course analysis of protein expression of 20°C pure methanol induction 112

Figure 3.6 Protein analysis of fermentation time course. Extractions were performed with sonication 113

Figure 3.7 Summary of unit operations involved in purifying and characterising expressed material 117

Figure 3.8 Qualitative protein data of purification sequence as described by Figure 3.7. 119

Figure 3.9 Mass balance of total host cell protein and total nucleic acids in the purification sequence described by Figure 3.7. 119

Figure 3.10 Electron microscopy image of k1,k1 Virus-Like Particles in a resuspended pellet after ammonium sulphate precipitation 120

Figure 3.11 Data for replicate fermentation attempt of *P. pastoris* Mut^S expressing tHBc-k1,k1 121

Figure 3.12 Qualitative protein data of time course analysis. 122

Figure 3.13 Close-up photo of 50mL tubes containing non-pelleted precipitate following ammonium sulphate treatment and centrifugation. 123

Figure 4.1 Schematic representation of a valve of a high-pressure homogeniser .. 128

Figure 4.2 Summary of unit operations involved in the analysis of buffer-related variables in cell disruption studies. 132

Figure 4.3 Effects of aggregation in clarified lysates after additions of various detergents at various concentrations. 134

Figure 4.4 High Pressure Homogenisation Response Surface Models 141

Figure 4.5 High Pressure Homogenisation performance optima summary. 143

Figure 4.6 Enzymatic cell disruption over time of cells incubated with lyticase 145

Figure 4.7 Total cell disruption of cells incubated with and without lyticase at various biomass concentrations. 146

Figure 4.8 Covaris components and parameters. 147

Figure 4.9 Total cell disruption through various methods 149

Figure 4.10 Qualitative dot blot analysis of crude and clarified lysates following lyticase treatment, AFA treatment and lyticase treatment followed by AFA treatment 150

Figure 4.11: Different tube types used in adaptive focused acoustics optimisation 153

Figure 4.12 Summary of Response Surface Models for Adaptive Focused Acoustic methods 154

Figure 4.13: Effect of vial design 155

Figure 4.14: Effect of micronised debris on purity. 159

Figure 4.15: Microscopy images of lysates 160

Figure 4.16: Microscale process mimicry. 162

Figure 5.1 Correlation between V_W and t when μ=0 171

Figure 5.2 Effect of methanol concentration on growth rates 173

Figure 5.3 Diagram of the methanol utilisation pathway in *P. pastoris* demonstrating nutrient competetion affecting heterologous protein expression. 185

Figure 5.4 Reaction diagram of the enzymatic conversion of a methyl ester on pectin to a carboxylated polymer and free methanol by pectin methyl esterase. 187

Figure 5.5 PNL-mediated depolymerisation of pection (left) to dGAME (right). 189

Figure 5.6 Metabolic analysis of *P. pastoris* Mut^S in several defined media conditions 192

Figure 5.7 Equilibirum between d-galacturonic acid and d-galacturonic acid methyl ester. 193

Figure 6.1 Overview of VLP production 200

Figure 6.2 Experimental overview. 203

Figure 6.3 Overview of methodology. 204

Figure 6.4 Time course profiles of residual methanol concentrations 207

Figure 6.5 Time course profiles of metabolic responses 209

Figure 6.6 Response Surface Model for biomass 210

Figure 6.7 Response Surface Model for Carbon Dioxide Evolution Rate 212

Figure 6.8 Response Surface Model for respiratory quotient 215

Figure 6.9 Response Surface Model for specific total core protein expression 217

Figure 6.10 Response Surface Model for specific soluble core protein expression 218

Figure 6.11 (A and B) Expression at various methanol flow rates of total and soluble tHBc respectively. (C and D) Expression at various glycerol flow rates of total and soluble tHBc respectively. (E and F) Expression at various induction temperatures of total and soluble tHBc respectively..... 221

Figure 6.12 Response Surface Model for specific VLP core protein expression..... 222

Figure 6.13 Correlation between specific soluble expression of tHBc and the relative proportion of soluble tHBc assembled into VLP-sized particles..... 223

Figure 6.14 Summary of response optima for methanol flow rate settings, glycerol flow rate settings and induction temperature settings..... 226

Figure 7.1 Representative metabolic anlaysis of different VLP constructs 233

Figure 7.2 Quantitative analysis of harvested material of different VLP constructs. 234

Figure 7.3 Metabolic analysis of HA2.3, $(M2e)_3$ constructs with different copy numbers at different induction modes. 236

Figure 7.4 Size Exclusion Chromatography fractions of bead-lysed HA2.3$(M2e)_3$ construct using pure methanol induction. 237

Figure 7.5 Size Exclusion Chromatography fractions of bead-lysedHA2.3$(M2e)_3$ construct using glycerol mixed feed induction. 238

Figure 7.6 Fermentation data of process scale-up of *P. pastoris* expressing tHBc-k1,k1 using pure methanol-based heterologous protein induction. 242

Figure 7.7 Fermentation data of process scale-up of *P. pastoris* expressing tHBc-HA2.3, $(M2e)_3$ using mixed feeding induction..... 244

Figure 7.8 Qualitative protein data of large-scale and small-scale fermentations, both using the same mixed feeding induction protocol 245

Figure 7.9 Carbon dioxide concentration profiles in fermentation exhaust gas and biomass concentration profiles for k1,LAH3, k1,$(M2e)_3$ and HA2.3,$(M2e)_3$ constructs during 30L mixed feed induction. 247

Figure 7.10 Qualitative analysis of expression of different tHBc constructs showing VLP-sized Tandem Core material through western blot analysis 248

Figure 7.11 Quantitative analysis of expression of different tHBc constructs showing VLP-sized Tandem Core material through dot blot analysis 249

Figure 10.1 Growth profile of master cell bank *Pichia pastoris* KM71H expressing tHBc-k1,k1. 292

Figure 10.2 Growth profile of working cell bank *Pichia pastoris* KM71H expressing tHBc-k1,k1. 292

Figure 10.3 Non-contaminated agar culture of working cell bank *Pichia pastoris* X33 expressing tHBc-GFP,e at 28°C, 1 week..... 293

Figure 10.4 Low-temperature induction fermentation data of *P. pastoris* Mut^+ expressing tHBc-GFP,e 294

Figure 10.5 Mixed feeding induction fermentation data of *P.pastoris* Mut^+ expressing tHBc-GFP,e 295

Figure 10.6 Fermentation data of *P. pastoris* Mut^{S} expressing tHBc-k1,k1 296

Figure 10.7 Total protein concentration of crude and clarified lysates of cells incubated in 1mL Yeastbuster^{TM} 298

Figure 10.8 Qualitative analysis of crude and clarified lysates following lyticase treatment and Yeastbuster^{TM}treatments 299

Figure 10.9 Comparison of lyticase-mediated cell disruption versus SoniPrep-mediated cell disruption with and without the addition of Triton-X100 after cell disruption.. 300

Figure 10.10 Reaction mechanism of d-galacturonic acid and methanol under acidic conditions 308

Figure 10.11 Response Surface Model for volumetric total core protein expression 309

Figure 10.12 Response Surface Model for volumetric soluble core protein expression 310

Figure 10.13 Time course profiles of volumetric tHBc expression 311

Figure 10.14 Response Surface Model for volumetric VLP core protein expression 312

Figure 10.15 Size Exclusion Chromatography fractions of bead-lysed k1, k1 construct using pure methanol induction 314

Figure 10.16 Log-scaled fermentation data of process scale-up of *Pichia pastoris* expressing tHBc-HA2.3, $(M2e)_3$ using mixed feeding induction. 315

List of tables

Table 4-1 Experimental overview for a 4-factor, two-level, full factorial experimental screening design with 3 midpoints to characterise homogenisation conditions. 135

Table 4-2 Relative factor contributions in a four-factor, two-level, half-fractional factorial screening designs 138

Table 4-3 Models and goodness of fit. 139

Table 4-4 Design space for a 3-factor, two-level, Central Composite Design with on-face axial points and three centre points. 141

Table 4-5 Relative factor contributions in a five factor, two-level, half fractional factorial screening designs. 151

Table 4-6 Relative factor contributions in a four-factor, two-level, full-fractional factorial screening design 152

Table 5-1: Relative factor contributions in a three-factor, two-level, full-fractional factorial screening designs 169

Table 5-2: Model and goodness of fit for specific growth rate (μ). 169

Table 5-3: Relative factor contributions in a three-factor, two-level, full-fractional factorial screening designs investigating induction of *P. pastoris* Mut^+ with pure methanol in microtitre plates. 176

Table 5-4: Similar to Table 5-3, this table shows the relative factor contributions of a screening design investigating induction of *P. pastoris* Mut^S with pure methanol in microtitre plates. 176

Table 5-5: Models and goodness of fit for screening responses investigating induction of *P. pastoris* Mut^+ with pure methanol in microtitre plates 176

Table 5-6: Similar to Table 5-5, this table shows the models and goodness of fit for screening responses investigating induction of *P. pastoris* Mut^S. 177

Table 5-7: Relative factor contributions in second three-factor, two-level, full-fractional factorial screening designs investigating induction of *P. pastoris* Mut^+ with pure methanol in microtitre plates. 181

Table 5-8: Similar to Table 5-7, this table shows the relative factor contributions of a screening design investigating induction of *P. pastoris* Mut^S with pure methanol in microtitre plates. 181

Table 5-9: Models and goodness of fit for second screening designs investigating induction of *P.pastoris* Mut^+ in microtitre plates. 181

Table 5-10: Similar to Table 5-9, this table shows the models and goodness of fit for screening responses investigating induction of *P. pastoris* Mut^S. 182

Table 6-1 Models and goodness of fit for final biomass concentration, Carbon Dioxide Evolution Rate, Respiratory Quotient, specific total expression of core protein, specific expression of soluble core protein, specific expression of VLP-sized core protein, volumetric total expression of core protein, volumetric

expression of soluble core protein and volumetric expression of VLP-sized core protein 205

Table 10-1 Fermentation media components 291

Table 10-2 Growth kinetics screening in microtitre plates 303

Table 10-3 Growth and expression kinetics screening in microtitre plates during methanol induction of *P. pastoris* Mut^{+} 304

Table 10-4 Growth and expression kinetics screening in microtitre plates during methanol induction of *P. pastoris* Mut^{S} 304

Table 10-5 Second screening of growth and expression kinetics screening in microtitre plates during methanol induction of *P. pastoris* Mut^{+} 305

Table 10-6 Second screening of growth and expression kinetics in microtitre plates during methanol induction of *P. pastoris* Mut^{S} 305

Table 10-7 Growth and expression kinetics screening in microtitre plates during PDM induction of *P. pastoris* Mut^{+} 306

Table 10-8 Growth and expression kinetics screening in microtitre plates during PDM induction of *P. pastoris* Mut^{S} 306

Table 10-9: Relative factor contributions in a two-factor, two-level, full-fractional factorial screening designs investigating induction of *P. pastoris* Mut^{+} in microtitre plates using defined media containing pectin derivatives. 306

Table 10-10 Relative factor contributions in a two-factor, two-level, full-fractional factorial screening designs investigating induction of *P. pastoris* Mut^S in microtitre plates using defined media containing pectin derivatives. 306

Table 10-11 dGAME Models and goodness of fit for second screening designs investigating induction of *P. pastoris* Mut^+ in microtitre plates using defined media containing pectin derivatives. 307

Table 10-12 dGAME Models and goodness of fit for second screening designs investigating induction of *P. pastoris* Mut^S in microtitre plates using defined media containing pectin derivatives. 307

Table 10-13: Central composite design for induction optimisation. Design space and corresponding response values. 313

List of symbols and abbreviations

[tHBc] or [TC-HBc]	Concentration of Tandem Core protein
[X]	Concentration of biomass
4PL	Four parameter logistic fitting
AFA	Adaptive Focused Acoustics
Ag	Antigen
ANOVA	Analysis of Variance
AOX	Alcohol Oxidase
APC	Antigen-Presenting Cell
BCA	Bicinchoninic Acid
BEI	Binary ethyleneimine
Beta-PL	Beta-propiolactone
BMGY	Buffered Glycerol-complex Medium
BODIPY	Boron-dipyrromethene
BSA	Bovine Serum Albumin
BSM	Basal Salts Medium
c-mol	Carbon moles
C^*	Oxygen concentration in gas phase
C_L	Oxygen concentration in liquid phase
CER	Carbon Evolution Rate
CHO	Chinese Hamster Ovary
cpb	Cycles per burst
CPP	Critical Process Parameter
DCW	Dry Cell Weight

DF	Duty Factor
dGA	d-Galacturonic Acid
dGAME	d-Galacturonic Acid Methyl Ester
DHAS	Dihydrosynthase
DO	Dissolved Oxygen
DoE	Design of Experiments
DOT	Dissolved Oxygen Tension
DSP	Downstream Processing
e	Empty insert
ELISA	Enzyme-Linked Immunosorbent Assay
EM	Electron Microscopy
FDA	Food and Drug Administration
GFP	Green Fluorescent Protein
GY	Glycerol
h	Hours
HA	Hemagglutinin
HBc	Hepatitis B core protein
HIC	Hydrophobic Interaction Chromatography
HPH	High Pressure Homogenisation
HPLC	High Performance Liquid Chromatography
I	Intensity
IL	Interleukin
kDa	kilo Daltons
kHz	kilohertz
k_La	Oxygen Mass Transfer Coefficient

LAH	Lower Alpha Helix
LPS	Lipopolysaccharide
LY	Lyticase
M1	Matrix protein 1
M2	Matrix protein 2
MCB	Master Cell Bank
MeOH	Methanol
MHC	Major Histocompatibility Complex
min	Minutes
MIR	Major Insertion Region
$mL/L_i/h$	Millilitres per litre per hour
N	Number of passes
NA	Neuraminidase
NP	Nucleocapsid Protein
OD	Optical Density
OTR	Oxygen Transfer Rate
OUR	Oxygen Uptake Rate
p	Pressure
P	Purity
PA	Polymerase Acidic protein
PAMP	Pathogen-Associated Molecular Pattern
PB	Polymerase Basic protein
PDM	Pectin Digest Medium
PRM	Pattern-Recognition Molecule
PRR	Pattern-Recognition Receptor

PS	Polysorbate
Q	Flow rate
R	Released protein concentration / Level of cell disruption
RdRp	RNA-dependent RNA polymerase
RFU	Relative Fluorescence Units
RMSE	Root Mean Square Error
RO	Reverse Osmosis
rpm	Revolutions per minute
RQ	Respiratory Quotient
R_s	Specific protein release
RSM	Response Surface Model
SEC	Size Exclusion Chromatography
T	Temperature
t	time
TCR	T-Cell Receptor
TEM	Transmission Electron Microscopy
tHBc or TC-HBc	Tandem Hepatitis B core protein
$tHBc_s$ or $TC\text{-}HBc_s$	Specific recovered Tandem Core protein
TLR	Toll-Like Receptor
TX	Triton-X
USP	Upstream Processing
V	Volume
VLP	Virus-Like Particle
Vol. or v/v	Volumetric
vRNP	Viral Ribonucleoprotein

V_W	Working Volume
WCB	Working Cell Bank
WCW	Wet Cell Weight
WHO	World Health Organisation
$Y_{sol., spec.}$	Specific soluble Tandem Core expression
$Y_{sol., vol}$	Volumetric soluble Tandem Core expression
$Y_{tot., spec.}$	Specific total Tandem Core expression
$Y_{tot., vol}$	Volumetric total Tandem Core expression
$Y_{VLP, spec.}$	Specific VLP expression
$Y_{tot., vol}$	Volumetric total Tandem Core expression
$Y_{VLP, spec.}$	Specific VLP expression
$Y_{VLP, vol.}$	Volumetric VLP expression
μ	Specific growth rate
vvm	Volume of air per volume of liquid per minute

Chapter 1 Introduction

1.1 Background

The Cough of Perinthus, described in 412 BC by Hippocrates, is possibly the first name attributed to what is now known as influenza or flu (Martin and Martin-Granel, 2006; Monto and Sellwood, 2013). Influenza has gained notoriety throughout history, with the Spanish Flu pandemic claiming more lives between 1918 and 1920 than the preceding Great War (Monto and Sellwood, 2013). Despite many advances made in the field of vaccine development, recent deadly outbreaks of flu variants serve as unfortunate reminders that pandemics of 1918 proportions cannot be ruled out in the future. Influenza virus remains a familiar threat to public health and continues to challenge the scientific community.

Most challenging is responding to the evolutionary adaptability of influenza virus, which often hinders effective prevention or treatment of influenza (Brooks *et al.*, 2013). Furthermore, current production methods are logistically and technologically inadequate to cope with pandemic surges, leaving a considerable number of individuals unprotected during such events.

This chapter aims to review current vaccine technologies and introduce the reader to the development of a universal influenza vaccine based on Tandem Core technology expressed in the yeast, *P. pastoris.*

1.2 Influenza Virology

1.2.1 Viral composition

Figure 1.1 shows a general structure of influenza virus. The virus is pleomorphic and is approximately 100-150 nm in diameter. Excluding influenza virus C, the influenza genome consists of 8 sections[1], encoding for ten genes[2] (Brooks *et al.*, 2013; Haaheim and Oxford, 2013).

The virus particle is surrounded by a lipid bilayer, derived from the host cell's plasma membrane and consisting of a variety of transmembrane proteins including hemagglutinin (HA), neuraminidase (NA), and matrix protein 2 (M2). Supporting the lipid bilayer is a matrix, composed of matrix 1 proteins (M1), that holds viral ribonucleoproteins (vRNPs). The vRNPs are composed of RNA, wrapped in nucleoprotein (NP) and relatively small amounts of nuclear export protein (NEP) also known as non-structural protein 2 (NS2). One type of vRNP is a polymerase complex composed of three polymerase proteins: polymerase basic protein 1 (PB1), polymerase basic protein 2 (PB2) and polymerase acidic protein (PA) (Samji, 2009; Brooks *et al.*, 2013; Haaheim and Oxford, 2013)

[1] Influenza C only possesses 7 RNA segments and hereby lacks the gene encoding for a key surface antigen: neuraminidase (Brooks *et al.*, 2013)

[2] If including a recently discovered gene encoding for the proapoptotic protein, PB1-F2, the total number of genes would be 11 (Chanturiya *et al.*, 2004; Samji, 2009) This gene is, however, absent in certain species of the virus, particularly those infecting swine (Chen *et al.*, 2001).

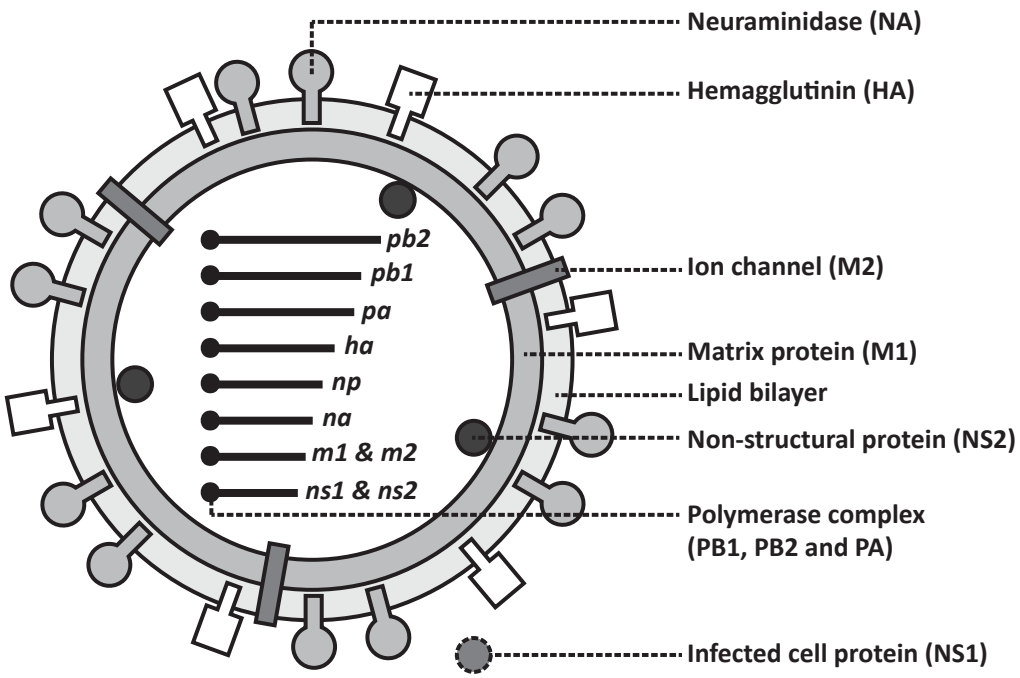

Figure 1.1 Schematic of influenza virus (Adapted from Virology Blog, 2009) [3]

1.3 Mechanism of infection

Figure 1.2 shows the lifecycle of influenza. After transmission of the virus into a host's body, the virus attaches itself to the surface of host cell, preferably to epithelial cells in the respiratory tract (Brooks *et al.*, 2013; Haaheim and Oxford, 2013) This attachment is mediated by HA, which binds to terminal sialic acids of glycolipids and glycoproteins on the cell surface. HA specificity depends largely on the type of linkage between sialic acid and the carbohydrates bound to glycoproteins or glycolipids. So-called α-2,6 linkage is generally associated with human influenza virus, whereas α-2,3 linkage is mostly associated with avian or equine influenza. Hemagglutinin associated with swine influenza can bind strongly

[3] Note that this is a general structure for influenza genera A and B, not C.

to both types of linkages (Skehel and Wiley, 2000; Samji, 2009; Haaheim and Oxford, 2013).

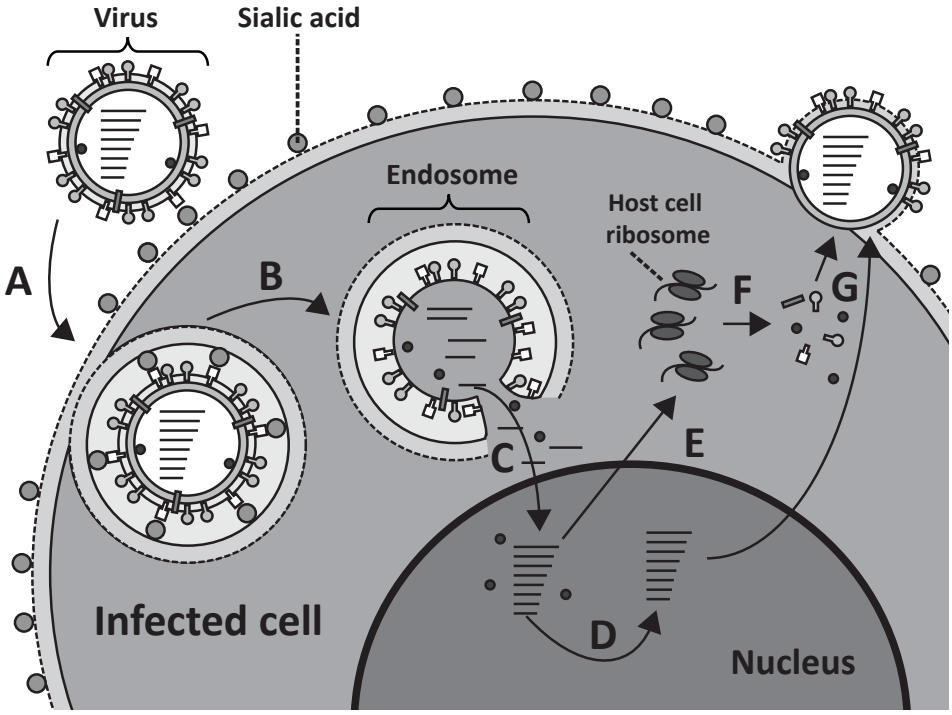

Figure 1.2 Influenza virus lifecycle. (A) HA binds to sialic acid leading to the formation of an endosome. (B) Fusion with endosomal membrane. (C) Transfer of viral RNA to host cell nucleus. (D) Replication of viral RNA. (E) Exportation of mRNA into the cytosol. (F) Translation of viral mRNA into viral proteins. (G) Assembly of new virus. (Adapted from Imperial College London, 2000)

Once the HA is bound to the cell surface, the cellular membrane encloses itself around the virus particle. This process, referred to as receptor-mediated endocytosis, creates a vesicle containing the virus particle inside the cell, called an endosome. The endosome allows entry of protons creating an acidic environment. This adjustment triggers a conformational change of HA, allowing it to fuse with the endosomal membrane, which exposes the virus matrix to the host cell cytoplasm. The M2 protein, which is an ion channel, also opens due to the acidic conditions within the endosome, allowing influx of protons, creating acidic conditions within

the virus particle. This results in the disassociation of vRNPs with M1 protein and ultimately the entry of vRNPs to the cytoplasm (Pinto, Holsinger and Lamb, 1992; Samji, 2009; Haaheim and Oxford, 2013).

The vRNPs travel to the host cell nucleus after they are released into the cytoplasm. Much remains unknown about the mechanism by which this happens. What is known is that some regions of the NPs in a vRNP function as Nuclear Localisation Signals (NLSs) that can bind to the cell's nuclear import cascade (Wu, Sun and Pante, 2007).

Once inside the cell nucleus, the vRNP complex uses its own enzymes to make positive sense cRNA copies of the negative sense vRNA template. Unlike most RNA polymerase, this particular viral polymerase does not require DNA as a template (Alberts *et al.*, 2004; Crow *et al.*, 2004). The vRNP instead uses its RNA-dependent RNA polymerase (RdRp) to produce positive sense RNA. The positive sense RNA can then either be used as a template to produce negative sense RNA for progeny virions or can be altered to function as messenger RNA to recruit virus proteins (Crow *et al.*, 2004; Haaheim and Oxford, 2013). In the latter case, for the RNA to function as messenger RNA, it needs to travel from the nucleus to the ribosomes in the cytoplasm to be translated to protein. Eukaryotic mRNA needs to be modified before it can leave the nucleus and therefore undergoes two steps prior to this: RNA capping and polyadenylation. RNA capping involves the addition of a guanine nucleotide with a methyl group at the 5' end of the mRNA transcript. Polyadenylation involves the trimming of a particular sequence of nucleotides at

the 3' end of the mRNA and subsequent addition of a so-called poly-A tail, which is a series of repeated adenine nucleotides (Alberts *et al.*, 2004). The viral genome of influenza mimics the poly-A tail, however, it does not contain an RNA cap (Samji, 2009). To circumvent this obstacle, the RdRp complex utilises a mechanism referred to as 'cap-snatching', which involves cleaving the 5' end of cellular mRNA and transferring this to the viral cRNA (Plotch, Bouloy and Krug, 1979). Once new negative sense vRNA strands have been produced from the cRNA templates and have formed vRNPs they are exported out of the nucleus via the Crm1 pathway. After sufficient viral genetic material and protein has been generated, the virus surface proteins (M2, HA and NA) assemble on the lumen of the host cell membrane. The other viral proteins assemble themselves beneath, through a yet unclear and debated process. The virus particles bud and are subsequently cleaved from the cellular membrane by NA (Samji, 2009; Haaheim and Oxford, 2013).

1.4 Viral taxonomy

Virus classification starts at the order taxon, followed by the family taxon, the genus taxon and, finally, the species taxon (Cann, 2012).

Influenza viruses are single-stranded ribonucleotide (ssRNA) viruses that are part of the family Orthomyxoviridae. This family does not belong to an assigned taxonomic order and consists of six genera (Cann, 2012; ICTV, 2012). Influenza virus is the collective name assigned to three genera of Orthomyxoviridae: influenza viruses A,

B and C (Cox and Hashimoto, 2011; Haaheim and Oxford, 2013). The first group can be divided into a range of species on which will be elaborated further on in this text.

1.4.1 Influenza genera

The first step in classifying a particular type of influenza virus is determining the genus (A, B or C).

As mentioned earlier, influenza virus C lacks the gene for neuraminidase, which is a major determinant of its genus. Further distinctions between genera can be made by antigenic variations in the matrix proteins (M1 and M2) and in the Nucleocapsid protein (NP). Because these genes are not cross-reactive, they are all determinants of the genus of influenza virus (A, B or C) (Brooks *et al.*, 2013). For human infection, influenza A and B are of greatest concern.

1.4.2 Species: antigenic drift and antigenic shift

To further understand the classification and nomenclature of influenza virus, particularly type A, one must consider the phenomenon of genetic reassortment. This can occur in two ways, called antigenic drift and antigenic shift.

Like most RNA viruses, the influenza virus has no genetic proofreading mechanism, causing as much as 10% of a virus population to be mutants. The mutations occurring in influenza virus often do not compromise the structure of proteins expressed. This is particularly true for the surface proteins HA and NA, where mutations mainly occur on the outer regions. These also happen to be the regions

that interact with antibodies upon infection of a host. Therefore, when sufficient changes occur in these regions, the adaptive immune system (1.5.2) might fail to neutralise the infectious agent, potentially leading to an epidemic outbreak. Such an accumulation of mutations in the genes coding for the antibody-binding sites is referred to as antigenic drift and can manifest itself as new field variants of influenza virus (Haaheim and Oxford, 2013).

More fundamental changes occur during another mode of genetic reassortment: antigenic shift. This can occur when two different influenza species co-infect the same host, as shown in Figure 1.3. Subsequently, during the assembly of progeny virions, the paternal RNA segments can mix. Particularly the crossing of HA and NA encoding genes can offer the virus a selective advantage because it can broaden the range of hosts the progeny virus can infect. This could be because different varieties of HA and NA have different affinities to the previously mentioned sialic acid linked receptors, which slightly differ per host type (Haaheim and Oxford, 2013).

Influenza virus B and influenza virus C do not have related viruses existing in the same hosts. This is why antigenic shift occurs exclusively in influenza virus A serotypes. Therefore, only influenza virus A has designated subtypes, based on antigenic differences exhibited by HA and NA (Brooks *et al.*, 2013). There are seventeen recorded subtypes of hemagglutinin (H1-H17) and ten subtypes of neuraminidase (N1-N10) (Tong *et al.*, 2012). Only four subtypes of hemagglutinin (H1, H2, H3, H5) and two subtypes of neuraminidase (N1, N2) have been recovered from humans (Brooks *et al.*, 2013).

1.4.3 Nomenclature

The nomenclature of a particular influenza virus serotype shows the following information in the following sequence: type (A, B or C), host animal of isolate (not for human host), geographic origin of isolation, strain number, year of isolation and – only in the case of influenza virus A - the particular antigenic variations of neuraminidase and hemagglutinin. Hence, an influenza virus A isolate from a person in Brisbane in 2007 can look like: *A/Brisbane/ 1/68(H3N2)*; an influenza virus B isolate can be named: *B/Brisbane/60/2008*; and an influenza virus A isolate from a pig can be labelled as: *A/swine/lowa/15/30(H1N1)* (Brooks *et al.*, 2013; Haaheim and Oxford, 2013).

1.4.4 Seasonal influenza

Annual seasonal epidemics of influenza occur due to the previously explained antigenic drifts in viral populations, causing slight amino acid sequence changes in HA and NA epitopes (Chowell, Nishiura and Bettencourt, 2007). This can result in the immune system not recognising the novel surface proteins and can cause previously acquired immunity to become ineffective (Nicholson, Wood and Zambon, 2003).

In the northern hemisphere influenza virus typically circulates from November to March. Influenza usually circulates between May and September in the southern hemisphere (Edlund *et al.*, 2011). However, in tropical regions, seasonal influenza occurs all year around and therefore, some argue, the term ‘inter-pandemic’

influenza is technically more appropriate (Lowen *et al.*, 2007; Van-Tam and Sellwood, 2013).

1.4.5 Pandemic influenza

An influenza pandemic is typically rare, in comparison to seasonal outbreaks of influenza. Such an event can be defined as: '*When a novel influenza A subtype spreads worldwide'* (Van-Tam and Sellwood, 2013). It should be stressed that seasonal influenza and pandemic influenza are not separate viral entities, but can rather be seen as different stages in the evolution of the virus. For instance, the A(H1N1) virus which emerged in 1918, caused the Spanish influenza pandemic lasting until 1920 (Monto and Sellwood, 2013). However, seasonal variants of the A(H1N1) virus kept circulating until 1956 (Van-Tam and Sellwood, 2013). These viruses were antigenic drifted version of the 1918 virus that had lost virulence, due to wide-scale increase of immunity among the general population.

Referring back to Van-Tam and Sellwood's definitions of the term pandemic, 'novel', in this context, is used to describe a type of influenza virus that the public has no or little pre-existing immunity against. Such a novel type of virus can arise via the previously explained phenomenon of antigenic shift. Wading birds appear to act as reservoirs for strains for influenza A and, occasionally, through close contact, strains can jump species directly. However, it is more common that a third species acts as an intermediary, as was seen in the 2009 swine flu pandemic (Monto and Sellwood, 2013). As illustrated by Figure 1.3, pigs can often act as mixing vessels for influenza A viruses, enabling antigenic shift (Nicholson, Wood and Zambon, 2003).

As previously discussed, this is because pigs have both α-2,6 linkage and α-2,3 linkage in the respiratory tract, making them susceptible to both avian and human influenza (Haaheim and Oxford, 2013).

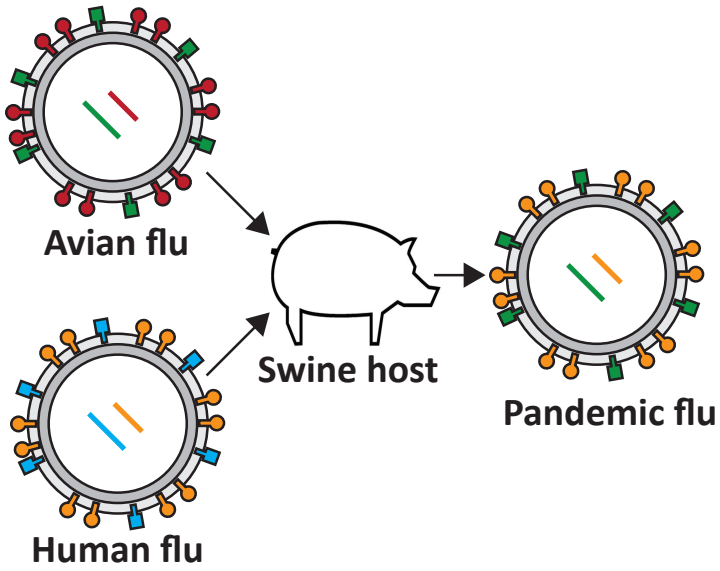

Figure 1.3 Antigenic shift in a swine host

1.5 Immunology

Immunity, defined as the active ability to resist disease, exists in two forms: innate immunity and adaptive immunity. Innate immunity, also known as natural immunity, is largely a function of an immediate response by phagocytes. This class of white blood cells includes natural killer cells, macrophages and granulocytes; which engulf and kill most pathogens. The activation of the innate immune response triggers the adaptive immune response (Madigan *et al.*, 2009).

Adaptive immunity is arguably more complex and has more relevance to the topic of vaccines. Unlike innate immunity, which is an immediate response, adaptive

immunity is acquired over time and its response can be up to 3 to 4 days after infection (Haaheim and Oxford, 2013).

1.5.1 The innate immune response

Phagocytes are the major constituents of the innate immune response and can recognise pathogens via membrane-bound Pattern-Recognition Molecules (PRMs) also known as Pattern-Recognition Receptors (PRRs) (Madigan *et al.*, 2009; Kreijtz, Fouchier and Rimmelzwaan, 2011). Such a molecule can interact with a so-called Pathogen-Associated Molecular Pattern (PAMP). For example, Toll-Like Receptor 4 (TLR-4), a human PRM, interacts with lipopolysaccharide (LPS), which is an endotoxin characteristic to gram-negative bacteria. The specific PRM-PAMP interaction causes a transmembrane signal which triggers the transcription of a variety of proteins that allow the digestion of the pathogen. The resulting smaller pieces of this digestion are called antigens. Once a phagocyte presents these on its outer membrane, the cell is called an Antigen-Presenting Cell (APC), which will then initiate the adaptive immune response. Note that not only phagocytic cells can be APCs. An APC is simply a cell displaying an antigen (Madigan *et al.*, 2009).

1.5.2 The adaptive immune response

A progenitor lymphoid cell can give rise to naïve T-cells and naïve B-cells. Naïve T-cells bear T-Cell Receptors (TCRs), which can bind to specific antigens; and two co-receptors: CD4 and CD8. A T-cell can differentiate to either a naïve CD4+ T-cell, which does not contain the CD8 receptor; or to a naïve CD8+ T-cell, which lacks the

CD4 receptor. If an endogenous antigen binds to the TCR of a naïve T-cell, this cell will be destroyed before it matures. The previously mentioned APCs can present antigens using one of two receptor classes: the class 1 Major Histocompatibility Complex proteins (MHC I); or the class 2 Major Histocompatibility Complex proteins (MHC II). MHC I receptors are found on all nucleated non-phagocytic cells, whereas MHC II receptors are only present on dedicated phagocytic APCs such as dendritic cells or macrophages. A naïve CD8+ T-cell can recognise foreign antigens bound to MHC I receptors causing the T-cell to activate to a cytotoxic-T cell. Cytotoxic-T cells destroy cells that are presented with the same antigen that caused their activation. A naïve CD4+ T-cell can recognise foreign antigens bound to MHC II receptors causing the T-cell to activate to a T-helper cell. T-helper cells enhance the immune response by secreting cytokines that activate phagocytes, natural killer (NK) cells and B-cells.

Naïve mature B-cells are coated in pre-made antibodies that can bind to antigens. A naïve mature B-cell initially functions as an APC by endocytosing and degrading these antigens, and presenting them on surface MHC II proteins to a subtype of T-helper cells, a T-helper 2 (Th2) cell. This type of T-cell can activate the B-cell by secreting cytokines, interleukin-4 (IL-4) and interleukin-5 (IL-5), causing the B-cell to differentiate into a memory B cell or a plasma cell. Memory B cells can live up to the order of years. When they reencounter the immunising agent that caused them to differentiate, they immediately turn into plasma cells without the assistance of T-cells. Plasma cells produce antigen specific antibodies. These antibodies can prevent the agent from infecting other cells by blocking membrane proteins, by

enhancement of phagocytosis (opsonisation) and by recruiting complement protein that causes damage to the infectious agent (Madigan *et al.*, 2009).

1.5.3 Immune response to influenza

It is thought that there are several PRMs involved in the innate response to influenza virus infection. Toll-like receptor 7 (TLR-7), present in mainly dendritic cells, recognises single-stranded viral RNA (Kreijtz, Fouchier and Rimmelzwaan, 2011). Retinoic acid inducible gene I (RIG-I) recognises 5'-triphosphates in genomics single stranded RNA, which appear after infection of the host cell (Pichlmair *et al.*, 2006; Pang and Iwasaki, 2011). Toll-like receptor 3 (TLR-3) bind double-stranded viral RNA. These interactions induce a cascade, resulting in the production of certain interferons and cytokines that inhibit protein production in host cells and, therefore, limit viral replication. They also stimulate dendritic cells and thereby promote antigen presentation to CD4+ and CD8+ T-cells, and ultimately viral clearance and immune memory (Kreijtz, Fouchier and Rimmelzwaan, 2011).

It is important to note that, although the immune system aims to protect the body from pathogens, paradoxically, it is often the cause of illness and death after exposure to these pathogens. For instance, it is thought the majority of deaths during the Spanish influenza pandemic were caused by an overactive immune response, manifesting in a so-called cytokine storm (Nicholson, Wood and Zambon, 2003).

1.5.4 Immunogenicity

There is a difference between antigens and so-called immunogens and the distinction between these terms is very relevant to the project. Antigens are substances that interact with TCRs or antibodies. This does not mean that this interaction always induces an immune response. It must therefore be stressed that an antigen that triggers an immune response is called an immunogen.

There are several intrinsic factors that determine whether an antigen is immunogenic.

The first is molecular size: generally, the larger an antigen, the more likely it is to be immunogenic. For example, haptens, which are low-molecular weight compounds, can interact with antibodies, but do not elicit an immune response. However, once these are coupled to a protein carrier, and thereby forming a larger molecule, these haptens become immunogenic.

Molecular complexity is another important indication of immunogenicity. Non-repeating polymers such as proteins and complex carbohydrates are often effective immunogens, unlike less complex molecules, such as lipids and simple polysaccharides (Madigan *et al.*, 2009).

Another important indicator of immunogenicity is the physical structure of the molecule. Generally insoluble macromolecules or protein aggregates elicit greater

immune response than the individual, soluble alternatives. This is because phagocytes are better at ingesting the former form.

There are also several extrinsic factors that contribute to immunogenicity. The dose of immunogen is a very important factor. Generally, doses between 10µg and 1g are effective. Anything outside this range may cause tolerance and can suppress a specific immune response. Another extrinsic factor is the ability of the immune system to recognise an antigen as foreign. Also, the route of administration has an effect on immunogenicity. For example, orally administered immunogens can degrade significantly before being taken up into the bloodstream (Madigan *et al.*, 2009).

1.6 Pharmaceutical interventions against influenza

1.6.1 Antiviral drugs

Pharmaceutical interventions are at hand if influenza infection cannot be prevented. There are two types of antiviral drugs targeted at influenza: M2 channel blockers (adamantanes); and neuraminidase inhibitors (Van-Tam and Lim, 2013).

M2 channel blockers prevent vRNA from entering a host cell. There are two types of channel blockers, both only active against influenza A viruses: Amantadine (Lysovir®, Symadine®) and Rimantadine (Flumadine®). Side effects of M2 channel blockers mainly involve the central nervous system, including anxiety, dizziness, concentration difficulty and insomnia (Sorrell *et al.*, 2011; Van-Tam and Lim, 2013)

Neuraminidase inhibitors prevent the release of progeny virions from a host cell. (Haaheim and Oxford, 2013). There are four NA inhibiting compounds: zanamivir (Relenza®), oseltamivir (Tamiflu®), peramivir (Rapiacta®, Peramiflu®) and laninamivir (Inavir®) (Van-Tam and Lim, 2013). The effectiveness of these antivirals can vary widely and is dependent on the strain of influenza itself. In the 2009 pandemic, antivirals such as zanamivir and oseltamivir were utilised in the early stages of the pandemic, but to little effect (Kawai *et al.*, 2009).

1.6.2 Vaccines

Vaccination is still considered the gold standard in protecting against influenza. There are various ways to prime the immune system for future influenza infections. All methods of influenza vaccination involve varying presentations of antigens and are listed below. For influenza, these antigens are almost exclusively HA, NA and M2 or derivatives of these surface proteins.

Seasonal influenza vaccines are commonly a trivalent dose, containing derivatives from three different strains of influenza A and B viruses to account for annual antigenic drift (Wolff and Reichl, 2008; WHO, 2009). However, the vaccine industry is in the process of moving to a quad-valent vaccine that will protect against the two most common A strains and two B strains of influenza. The standard dosage of vaccines is typically 15µg per strain (CDC, 2018).

1.6.2.1 Live, attenuated virus vaccines

Modifying or reassortment of a 'wild type' pathologic virus, also known as reverse genetics, produces a class of vaccines, referred to as live attenuated virus vaccines (Neumann, Ozawa and Kawaoka, 2012). The result is a virus that maintains its ability to replicate and cause an immune response however, usually, without causing disease. In the case that disease is caused by an attenuated-type vaccine, symptoms are weaker and are generally referred to as adverse reactions (CDC, 2011). Attenuated viruses generate the strongest and longest acting immune response out of all vaccine classes. However, they are unsuitable for immunocompromised individuals and carry the theoretical risk of pathogenic reversion. Also, it is very difficult to standardise, maintain and select certain strains (Madigan *et al.*, 2009). This is particularly difficult with influenza virus, due to its high mutability.

MedImmune have licensed a live attenuated vaccine that can only replicate at 25°C. This cold-adapted virus is therefore isolated in the nasopharynx and cannot spread to the lungs (EMA, 2013).

1.6.2.2 Inactivated influenza vaccines

This class of vaccine involves the expansion of the pathogen in question in a particular host. In the case of influenza virus vaccine production this is achieved in the allantoic fluid of fertilised chicken eggs. After harvesting this fluid, the virus is inactivated, disabling replication. This is typically achieved using detergents. The virus fragments recovered from this process are subsequently treated to ensure any genetic material has been removed. The resulting inactivated vaccine is specifically

referred to as a split-virion vaccine (Nakayama *et al.*, 2012). If the former fraction is further purified to obtain a higher concentration of antigen (typically HA and NA), by removing more internal viral proteins, the resulting vaccine is referred to as a subunit vaccine (College of Physicians of Philadelphia, 2014; Talbot *et al.*, 2015). Overall, inactivated virus vaccines generally yield relatively short-lived immune responses (Madigan *et al.*, 2009).

1.6.2.3 Virus-like particle vaccines

Virus-like particles (VLPs) are structures consisting of multiple proteins that are arranged to collectively mimic authentic native viruses. Although a VLP typically contains nucleotides originating from its host expression system (i.e. a genetically engineered microbe), a VLP is not considered infectious because it lacks *viral* genomic material (Roldão *et al.*, 2010).

Compared to individual viral peptides or proteins, VLPs present epitopes that have more conformational similarity to the native virus. As these multiprotein structures mimic the conformation and organisation of authentic native viruses, the immune response is expected to be higher than for individual peptides or proteins (Roldao *et al.*, 2010). Most vaccine classes are formulated with adjuvants to enhance immune response. VLPs usually have highly repetitive surfaces, which allow cross-linking on specific B-cell receptors. VLPs can therefore induce strong B-cell responses, without requiring adjuvant (Jegerlehner *et al.*, 2002; Roldão *et al.*, 2010). VLPs can also be employed as a scaffold to bear one or several antigens. This can enable broader and more efficient protection. The potency of such VLPs triggers relatively high immune

responses, while using lower amounts of antigen, potentially reducing vaccine costs significantly (Roldão *et al.*, 2010).

1.6.3 Adjuvants

Some antigens, especially smaller targets, can be poorly immunogenic and require further assistance to achieve the right levels of immuno-stimulation. Thus, adjuvants are often added to the formulated vaccine product. Adjuvants are substances that cause a non-specific immune enhancement, making the immune system more receptive to the accompanying vaccine (National Research Council Canada, 2010). Adjuvants come in many forms, from bacterial toxins to non-biological agents such as aluminium hydroxide (Glaxo Smith Kline, 2005).

1.7 Vaccine production

Once candidate viral strains has been selected based on recommendations from the World Health Organisation (WHO), candidate high-growth seed strains are distributed to various collaborating centres, depending of the region (Gerdil, 2003). Manufacturers can use these recommendations to prepare a range of vaccines. A selection of methods of vaccine production is described below.

1.7.1 Fertilised chicken egg-derived vaccine production

The most widely used method for vaccine production involves injecting live influenza viruses into the amniotic sack of a nine to twelve-day old fertilised egg (WHO, 2009; Cox and Hashimoto, 2011). Each egg typically yields one to two doses

of vaccine (Layton and Lenfestey, 2005b). After an incubation period of up to 72 hours, the virus or virus derivative is harvested (Tree *et al.*, 2001; Roland, 2014). Eggs are generally used to produce live attenuated vaccines or inactivated virus vaccines (Palache, Brands and van Scharrenburg, 1997; Wolff and Reichl, 2008).

1.7.1.1 Bioprocessing

The following describes one possible method to purify egg-derived vaccines. The allantoic fluid of the embryo is harvested and is subsequently clarified by centrifugation. Using tangential flow ultrafiltration, the volume of the supernatant is reduced. The retained virus is then resuspended overnight in Ca-Mg-saline solution. This suspension is subsequently loaded onto a sucrose gradient and centrifuged, which allows the virus to be pooled and collected (Liu *et al.*, 1995). There are, however, a number of ways to purify a vaccine. For instance, on larger scales, the latter step might not always be feasible and can be substituted by a variety of chromatographic steps (Wolff and Reichl, 2008). Figure 1.4 outlines the methods to which various downstream objectives can be achieved.

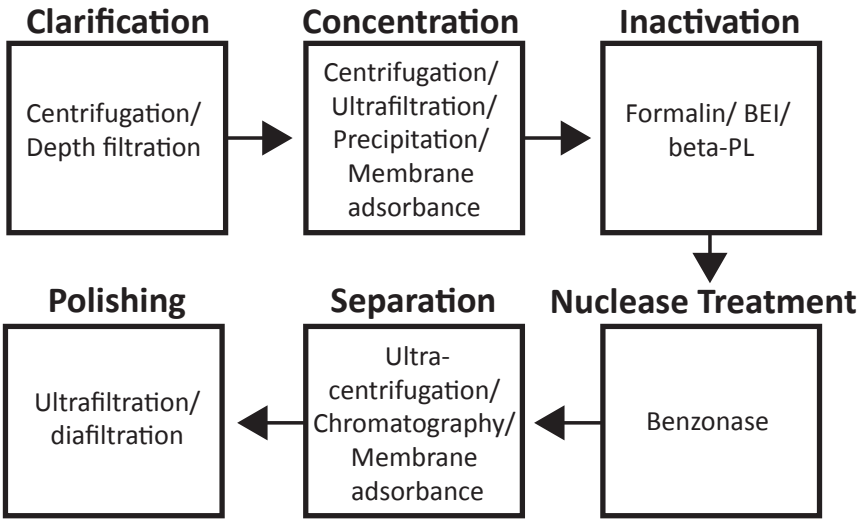

Figure 1.4 Downstream process of egg and mammalian cell derived vaccines (adapted from Wolff and Reichl, 2008).

1.7.1.2 Advantages and disadvantages

Vaccines have been produced in fertilised eggs for a long time. It is therefore a well-established and widely approved process.

However, there are a significant number of disadvantages to this particular production platform. Due to the general requirement of vaccine tri-valency or quad-valency; multiple eggs are required for a single dose of vaccine. Therefore, prior to vaccine production, a considerable number of fertilised eggs need to be supplied. The supply of such an amount of eggs can take up to six months and poses a significant limit to the scalability of vaccine production (Gerdil, 2003). Furthermore, post-translation mechanisms can alter the composition of wild type human HA expressed in hens' eggs (Oxford *et al.*, 1987, 1991; Robertson *et al.*, 1987) leading to relatively lower immune responses and decreased efficacy of vaccines produced in this fashion (Wood *et al.*, 1989; Tree *et al.*, 2001). Additionally, residual egg albumin in egg-derived vaccines is known to cause allergic reactions in some individuals

(Nayak, Lehmann and Reichl, 2005). A significant disadvantage is that this production method can be highly unsuitable for avian related influenza A vaccine production. For instance, the avian related H5N1 influenza A strain of 1997 showed a mortality rate of 70%-100% in chickens (Cox and Hashimoto, 2011). Chicken embryos used for the production of a vaccine against this virus were therefore killed before significant amounts of vaccine could be produced.

Much of the industry has established capacity in egg-based vaccine production. Economic drivers are not yet present to warrant a shift from this established process. The reasons for this are a combination of high costs and risk; and low profit margins: The majority of costs of developing and producing a new vaccine will be dedicated to research and development; and demonstrating safety and efficacy in clinical trials. Secondly, parties who buy the vaccine, predominantly governments, typically offer a buying price that only covers the cost of producing a single dose of vaccine. For instance, in the 2004-2005 influenza season in the United States, the highest recorded wholesale price of influenza vaccines was USD 8.50 (Layton and Lenfestey, 2005a).

1.7.2 Vaccine production in mammalian cells

As with fertilised eggs, mammalian cells are generally used to produce live attenuated vaccines, inactivated virus vaccines or split virion vaccines (Palache, Brands and van Scharrenburg, 1997; Wolff and Reichl, 2008). There are, however, instances where subunit and VLP vaccines are produced using this expression

system (Szymczakiewicz‐Multanowska *et al.*, 2009; Wu *et al.*, 2010). A variety of mammalian cell lines can be used for the production of vaccines. Popular expression systems are Madin-Darby canine kidney (MDCK) cells and African green monkey kidney (Vero) cells, due to their ability to produce high virus titres (Merten *et al.*, 1996; Tree *et al.*, 2001).

1.7.2.1 Bioprocessing

Mammalian cells can be cultured in serum or serum-free medium. As with eggs, the cells are infected with virus. HA screening can be employed to monitor virus activity during growth and the number of live viruses can be measured by solid plaque assays (Mahy, 1985; Tree *et al.*, 2001). Once the required amount of virus production has been determined and cultures have been appropriately diluted, the cultures are further expanded by inoculating, for example, roller bottles or spinner flasks containing micro-carriers. For larger scale production, micro-carriers can be used (Genzel *et al.*, 2004; Genzel, Fischer and Reichl, 2006; Chen *et al.*, 2010).

The downstream processing of mammalian cell-derived vaccines is very similar to that of fertilised egg-derived vaccines (Figure 1.4).

1.7.2.2 Advantages and disadvantages

Vaccine production in mammalian cells offers numerous advantages to vaccine production in fertilised eggs. Mammalian cell-derived vaccines have been shown to trigger similar immune responses as naturally circulating strains (Nerome *et al.*, 1999) and have shown to be safe and efficacious in clinical studies (Palache *et al.*,

1999; Percheson *et al.*, 1999). Furthermore, mammalian cell-derived vaccine production is scalable and can employ the use of well-characterised and stable cell lines; simplifying the logistics of production in comparison to the use of eggs. Therefore, processing time for mammalian cell-derived vaccines is significantly less than that of egg-derived vaccine production.

Generally, the use of live influenza viruses is required in this expression system and poses significant risks with respect to biological containment. Also, the risk of contamination by pathological agents is increased when the use of animal derived products, such as serum, is required (Nims, 2006; Chen *et al.*, 2008). Furthermore, the productivity of the process depends heavily on the cell line employed. In addition, a reassortant virus, which contains genes from two parental viruses, often needs be employed to achieve high productivity (Tree *et al.*, 2001; Cox and Hashimoto, 2011; Ellis, Rappuoli and Ahmed, 2013).

1.7.3 Vaccine production in *Escherichia coli*

E. coli is a popular expression system for a variety of recombinant proteins. Influenza viruses cannot use this prokaryotic organism as a host and therefore *E. coli*-derived vaccines are limited to those of the subunit and VLP classes.

One current method being developed for the production of influenza vaccines is the production of M2e fusion proteins in *E. coli* (Fiers *et al.*, 2009; Kim *et al.*, 2013; Deng *et al.*, 2015). M2e is the outer domain of the M2 protein and is 23 amino acids long. The M2 ectodomain is highly conserved and cross-reactive, making it a good

candidate for a universal vaccine. However, low immune responses to this particular antigen not exerting any evolutionary pressure on the sequence could be the cause of this high conservation. Because of this low immune response, M2e has to be fused with more immunogenic antigens such as VLPs, for example Hepatitis B core protein (HBc), formed by the Hepatitis B virus (Fiers *et al.*, 2009; Deng *et al.*, 2015); or flagellin (Huleatt *et al.*, 2008).

1.7.3.1 Bioprocessing

The expression of M2e-flagellin fusion protein *in E. coli*, described by Huleatt *et al.* (2008) will be discussed to illustrate upstream production of microbial-derived recombinant vaccines.

Initially, a DNA construct is created consisting of 4 tandem copies of M2e fused to the *Salmonella typhimurium* flagellin fljB (STF2) gene. The resulting plasmid is then inserted into competent *E. coli* cells. The transformed *E. coli* cells are subsequently expanded through large-scale fermentation, which is a well-established process (Hewitt *et al.*, 2000; Huleatt *et al.*, 2008).

At the end of the fermentation process, the cells are harvested through centrifugation. The sediment is resuspended and lysed in buffer containing lysozyme and a protease inhibitor cocktail. This mixture will facilitate lysis while minimising proteolytic cleavage of the product. Following lysis, the suspension is centrifuged to separate cell debris and other insoluble materials. The supernatant is subsequently passed through a depth filter to protect chromatographic columns

further downstream. The first chromatographic step is a gradient ion exchange separation that captures the product. To remove endotoxin, Size Exclusion Chromatography is used. The final stages of downstream processing include ultrafiltration and sterile filtration for formulation (Huleatt *et al.*, 2008; Wolff and Reichl, 2008; Josefsberg and Buckland, 2012). The process is summarised in Figure 1.5.

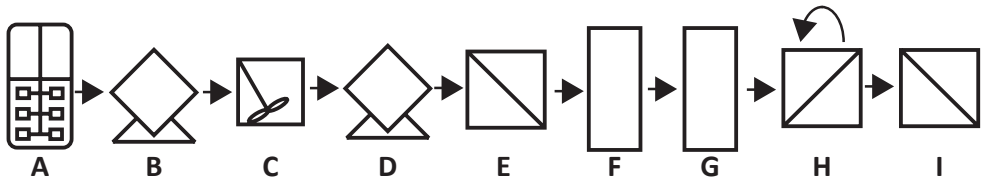

Figure 1.5 Schematic overview of the production process for the manufacturing of M2e fusion proteins. (A) Fermentation. (B) Harvest centrifugation. (C) Cell lysis. (D) Clarification centrifugation. (E) Depth filtration. (F) Gradient ion exchange chromatography. (G) Size Exclusion Chromatography. (H) Diafiltration. (I) Sterile filtration.

1.7.3.2 Advantages and disadvantages

E. coli is one of the most characterised organisms in the field of biochemistry. Also, in comparison to any other expression system, *E. coli* grows fast and is highly productive. These particular characteristics make *E. coli* an excellent candidate for vaccine production in the event of a pandemic surge. Furthermore, *E. coli* is easily transformed and upstream production is relatively cost-effective (Verma, Boleti and George, 1998).

However, recombinant type vaccine production in *E. coli* can also result in unpredictable clinical efficacy and safety. It has been reported that after the vaccination of pigs with a fusion protein of M2e and influenza nucleoprotein (M2e-NP), the vaccine aggravated influenza symptoms rather than preventing them

(Heinen *et al.*, 2002). Furthermore, *E. coli* produces endotoxin, which can cause inflammatory reactions if not removed from a vaccine. Finally, *E. coli* can be quite restrictive as an expression system for more complex recombinant mammalian proteins because it does not possess the post-translational modification pathways required (Verma, Boleti and George, 1998). To date, no licensed influenza vaccine has been produced in *E. coli.*

1.7.4 Vaccine production in insect cells

On October 22 2013 Protein Sciences Corporation's "FluBlok" seasonal influenza vaccine was FDA approved and is currently marketed throughout the U.S.(FDA, 2013; Protein Sciences Corporation, 2013). It is the first influenza vaccine produced in insect cells.

1.7.4.1 Bioprocessing

The vaccine consists of trivalent recombinant hemagglutinin (rHA) proteins produced in *Spodoptera frugiperda*. This is one of the most popular insect host used in industry for the production of recombinant proteins (Summers, Smith and Station., 1987). The rHA is expressed in the insect cells through infection with baculovirus, a popular insect expression system (Verma, Boleti and George, 1998; Cox and Hashimoto, 2011). First a HA gene is cloned using RT-PCR resulting in RNA being transcribed to cDNA. The DNA is subsequently introduced into a baculovirus vector developed by Protein Sciences Corporation (Protein Sciences Corporation, 2013). Using calcium phosphate precipitation, the vector is inserted into *Spodoptera*

frugiperda Sf9 cells together with linearised genomic *Autographa californica* nucleopolyhedrovirus (AcMNPV) DNA to enable homologous recombination.

After fermentation, the infected cells are harvested and separated from the medium using disk stack centrifugation. After the pellet is collected, it is exposed to a non-ionic detergent in order to dissolve the insect membranes. A depth filter is subsequently used to remove cell debris. This is followed by three chromatographic processes. The first is a cation exchange process (CEX), which captures the vaccine product and other protein contaminants. The second process is Hydrophobic Interaction Chromatography (HIC), which retains rHA and discards protein contaminants. The third and final chromatography step is a Q-membrane, which removes nucleotides. Finally, the eluate passes through an ultrafiltration module for formulation. The process is summarised in Figure 1.6.

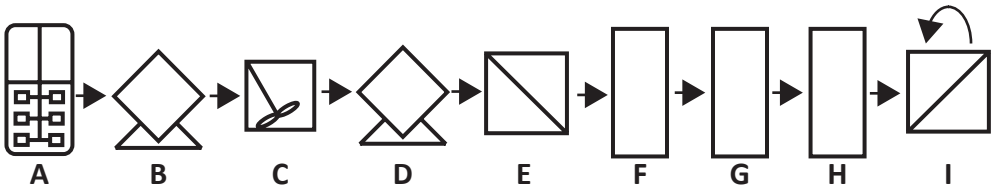

Figure 1.6 Schematic overview of the production process for the manufacturing of rHA (adapted from Blaha *et al.*, 2012). (A) Fermentation. (B) Harvest centrifugation. (C) Cell lysis. (D) Clarification centrifugation. (E) Depth filtration. (F) Cation exchange chromatography. (G) Hydrogen interaction chromatography. (H) Size Exclusion Chromatography. (I) Diafiltration.

1.7.4.2 Advantages and disadvantages

The risk of contamination by advantageous agents in insect cell cultures is significantly lower than in mammalian cell cultures because insect cells do not require any animal-derived products, such as serum, for growth (Nims, 2006; Chen

et al., 2008). Furthermore, the chance of insect cells becoming hosts for vertebrate viruses is relatively small, because of the evolutionary distance between mammals and insects (Kuno and Chang, 2005). The only viruses known to propagate in both mammals and insect cells are arboviruses. This virus does not pose a serious risk to the production of vaccines using this expression system because the susceptibility of *Spodoptera frugiperda* Sf9 to arboviruses is very low (Cox and Hashimoto, 2011). The only known arbovirus to cause a significant infection is the mosquito-borne St. Louis encephalitis virus (Verma, Boleti and George, 1998). An advantage that the arthropod expression system has over prokaryotic expression systems, such as *E. coli*, is that insect cells are able to perform most eukaryotic post-translational alterations, generally resulting in higher clinical efficacy. The polyhedrin promoter used in insect-cell cultures is considerably stronger than most eukaryotic promoters. Therefore, this type of expression system produces higher levels of protein than mammalian cell cultures (Luckow and Summers, 1989).

For the purposes of human vaccine production, insect cells have more appropriate post-translational mechanisms than *E. coli*. However, insect cells process mature oligosaccharides differently to mammalian cells and do not recognise certain proteolytic cleavage sites on mammalian proteins (Hsieh and Robbins, 1984; Kuroda *et al.*, 1986). These drawbacks could result in lower clinical efficacy of insect-derived vaccines in comparison to mammalian-derived vaccines.

1.8 Tandem Core technology

This thesis will focus on the development of a production process for a novel *universal* influenza A vaccine. *Universal* implies a vaccine that is reactive with multiple types of influenza A, accounting for annual antigenic drift and unpredictable antigenic shift.

A strategy to increase cross-protection against variants of influenza virus is to develop a vaccine consisting of highly evolutionary conserved epitopes of influenza virus. Examples of such epitopes are the aforementioned M2e peptide, influenza nucleoprotein (NP) and the stem of hemagglutinin (Fiers *et al.*, 2009; Abate, 2013; Kim *et al.*, 2013; IQur, 2014). However, the absence of evolutionary pressure on these, usually structural, epitopes is likely due to a lack of recognition by the immune system. Thus, it is often found that such epitopes alone elicit no significant immune response. However, the immunogenicity of these poorly immunogenic antigens can be increased through presentation on the surface of a Virus-Like Particle.

Compared to individual viral peptides or proteins, such as M2e, VLPs present epitopes in such a way as to achieve higher conformational similarity to a native virus.

Hepatitis B Core protein (HBc), a monomer that self-assembles into VLPs, has been a popular carrier of a variety of heterologous epitopes derived from many pathogens (Roose *et al.*, 2013). This fusion-facilitated versatility of HBc allows for

the creation of multivalent VLP-based vaccines, i.e. vaccines consisting of VLPs that present a variety of heterologous epitopes. Such a strategy would also be beneficial for the development of a universal flu vaccine as this would enable higher cross-protection than employing single epitopes or monovalent vaccines. However, it is very challenging to control and determine the amount of each heterologous epitope present in an assembled VLP for a multivalent vaccine.

A way to control the amount of individual heterologous epitopes present in a single VLP is to attach a flexible linker between two HBc monomers Figure 1.7. In this way, attaching a different heterologous epitope to the major insertion region (MIR) of each monomer, results in the expression and subsequent assembly of VLPs presenting an equal amount of each epitope. The linking of these two HBc monomers, via use of a flexible linker, will be referred to as Tandem Core technology (iQur, 2014).

The formation of Tandem Core-based VLPs occurs through the assembly of intermediates consisting of two, three and five fused Tandem Core dimers (Holmes *et al.*, 2015)

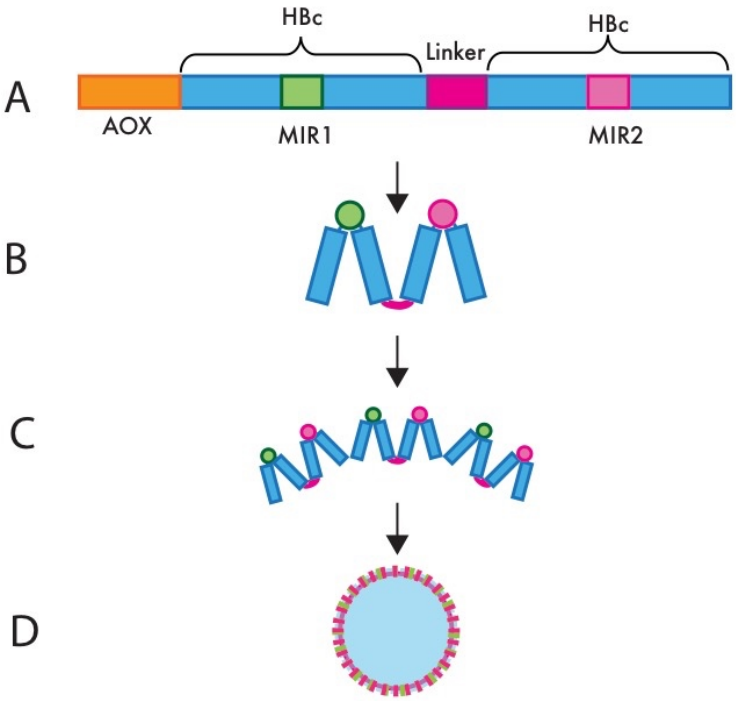

Figure 1.7 Tandem Core Technology overview: (A) DNA sequence of construct. AOX promotor followed by two HBc units connected by a flexible linker sequence. Within each HBc sequence is a Major Immunodominant Region (MIR) sequence, allowing for the insertion of heterologous sequences of interest. (B) Individual tandem Hepatitis B core (tHBc) unit. (C) Self-assembly of tHBc units into a VLP (D).

Tandem Core technology has several advantages over conventional VLPs. Firstly, it can hold large twin inserts, whereas conventional VLPs can only be fused to single and small peptides. Furthermore, Tandem Core vaccines will not require the use of adjuvants. Finally, Tandem Core VLPs can act as more agile antigen presentation scaffolds than conventional VLPs, allowing for the combination of multiple epitopes associated with a wide variety of diseases. This property makes Tandem Core technology a good candidate central component for a universal vaccine production platform. Even though Tandem Core technology could be applied for the presentation of epitopes associated with a wide array of diseases, the scope of this thesis will be restricted to the development of a universal influenza vaccine.

1.9 Pichia pastoris

Tandem Core VLPs have another important advantage over conventional, egg-derived vaccines. VLPs, in general, can be produced in suspension cultures, offering the benefit of scalability.

VLPs have previously been successfully produced in suspension cultures using *P. pastoris* (Freivalds *et al.*, 2011; Jiang *et al.*, 2011). Indeed, in this series of experiments, a variety of Tandem Core VLPs will be expressed by *P. pastoris.*

P. pastoris is a single-celled yeast and, since yeasts are eukaryotic, it is more likely that proteins typically expressed in a eukaryotic environment are appropriately translated and assembled in *P. pastoris* than in non-eukaryotic expression systems such as *E. coli*. However, the most important reason for choosing to express Tandem Core VLPs in *P. pastoris* rather than in *E. coli*, is the absence of lipopolysaccharides (LPS) in the purified product (iQur, 2014).

What distinguishes *P. pastoris* from many other yeast species, like *Saccharomyces cerevisiae*, is its preference for respiratory growth, i.e. its use of molecular oxygen as a terminal electron acceptor during the synthesis of ATP, rather than fermentative growth. A great advantage of this is that the carbon source is directly turned to biomass, rather than to toxic by-products, such as ethanol. This results in high-density cell cultures (Cregg, 2007). It is worthwhile to note that despite this, a general challenge associated with *P. pastoris* is low yields of recombinant protein

when compared to other expression systems, such as *E. coli*; insect cells; and Chinese Hamster Ovary (CHO) cells (Weinacker *et al.*, 2013; Maccani *et al.*, 2014).

P. pastoris can use methanol as a sole source of carbon and because of this characteristic, a specific set of genes can be induced to produce the following metabolic enzymes: alcohol oxidase (AOX) and dihydrosynthase (DHAS). The regulatory genes involved in the transcription, *aox1, aox2* and *das* respectively, constitute a major advantage of the *P.pastoris* expression system as they are responsible for approximately 30% of protein production in methanol-induced cultures, with *AOX1* accounting for the majority of protein production (Couderc and Baratti, 1980; Tschopp *et al.*, 1987; Krainer *et al.*, 2012).

P. pastoris also differs from other expression systems, such as *E. coli*, in that vectors used for genetic recombination are integrated into the host genome. This happens via homologous recombination at sites where there is significant homology between sequences (>0.5 kb) (Cregg, 2007). For *P. pastoris* transformations, these sites are typically two aligned AOX1 promoter (pAOX1) sequences (Cregg *et al.*, 1985).

1.9.1 Bioprocessing – fermentation

As shown in Figure 1.8, a typical bioreactor-based fermentation of *P. pastoris* consist of three phases: The batch phase, a glycerol fed-batch phase and the induction phase, usually using methanol.

The objective of the first stage, the batch phase, is to maximise culture biomass through the consumption of glycerol. As the culture grows, more oxygen is required to support growth and cell maintenance. It is therefore important that the Dissolved Oxygen (DO) content is controlled and maintained above 20%. This is typically achieved through a cascade of increased culture agitation followed by gas blending of pure oxygen in the sparged gas. The batch culture is grown until the glycerol in the medium has been completely consumed. This stage is usually reached after 18-24 hours and will be indicated by a rapid increase in DO, also known as a DO 'spike', or a rapid decrease in the Carbon Evolution Rate (CER).

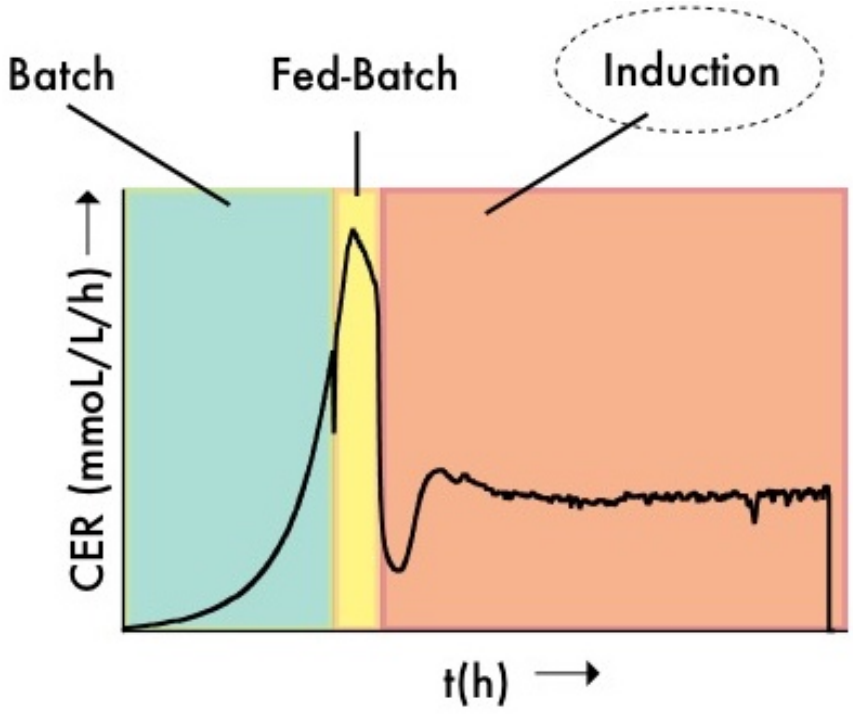

Figure 1.8: Typical fermentation of *P. pastoris*. Carbon Evolution Rate (CER) is shown over time, t. From left to right: Batch phase, fed-batch phase and induction phase.

Once glycerol has been completely consumed during the batch phase, a carbon-limiting glycerol feed is introduced for at least 4 hours. This stage serves two purposes: firstly, it allows for control in achieving a target biomass; secondly, it

gradually creates a carbon-limiting environment to facilitate adaptation to methanol during the induction phase of fermentation.

During this final stage, heterologous protein is produced by the cells. The start of the induction phase involves the adaptation of the cell culture to a gradually increasing feed of methanol as a sole carbon source. This gradual feeding approach is preferable, as an immediate addition of high volumes of methanol would poison the culture. This adaptation period is typically a few hours, while the total induction phase can range between 24 and 100 hours in duration (Invitrogen Corporation, 2002; Mirro, 2011). Tandem Core technology, expressed in *P. pastoris,* offers an exciting proposition for creating a future universal influenza vaccine. However, as with all novel concepts, development and characterisation of applying this technology will be required and process methodologies should be developed accordingly.

1.10 Thesis aims

A wide variety of methods exist for producing influenza vaccines. However, these platforms cannot fulfil public demand in the event of a pandemic due to limitations in scalability, nor do these produce efficacious vaccines accounting for antigenic drift and antigenic shift.

This research, which represents a component of research performed by an FP7[4] consortium project, will focus on using Tandem Core technology to develop a universal influenza vaccine. Using this technology, recombinant Virus-Like Particles can be expressed in bioreactor-based suspension cultures of *P. pastoris*, providing a highly scalable production platform. To realise the potential of Tandem Core technology as a central component in a universal vaccine production platform, the production of Tandem Core VLPs in bioreactor suspension cultures needs to be extensively characterised.

The specific objectives of this thesis are therefore:

Initial process characterisation of novel Tandem Core VLP constructs (Chapter 3)

This chapter will aim to identify critical process parameters associated with upstream and downstream processing of novel Tandem Core VLP constructs. This will involve investigating the effects of changing constructs, yeast phenotype, fermentation modes and methods of primary recovery.

Characterisation, miniaturisation and optimisation of primary recovery methods (Chapter 4)

The utilisation of scale-down tools to increase process throughput is a key feature in this thesis however, as the product is intracellular, appropriate, low sample volume, lysis methods need to be developed. This chapter will describe research aimed to

[4] European Union's Seventh Framework Program for Research and Technological Development

address the effects of various primary recovery methods, with an emphasis on cell disruption, on the yield of soluble Tandem Core material. In doing so, the chapter will discuss how cell disruption processes are characterised, optimised and miniaturised.

Microscale upstream process platform development (Chapter 5)

Having developed the prerequisite tools to implement in miniaturised upstream process development, the use of miniaturised fermentation platforms can be investigated to address several key observations made in chapter 3, and to identify appropriate scale-down platforms and expression techniques for monitoring product expression. This chapter will therefore aim to develop a miniaturised fermentation platform intended for rapid upstream process characterisation.

Characterisation of Tandem Core Virus-Like Particle production (Chapter 6)

This chapter aims to use miniature bioreactors, DoE-based methodology, and the aforementioned scale-down cell disruption tools, to study Virus-Like Particle production in a bioprocessing context through variance of previous identified critical process parameters and windows of operation, to maximise VLP expression.

Scale-up and application of Tandem Core Technology to universal influenza vaccine candidates (Chapter 7)

The previous chapter will have addressed the characterisation of VLP production at miniaturised scale using a simplified, non-epitope-exposing, variant of Tandem Core VLP. This chapter aims to translate these findings to industrially relevant vaccine production scenarios by investigating the effects of scale-up and variation of epitope insets that could be utilised as a universal influenza vaccine.

Chapter 2 Materials and methods

This chapter describes common materials, routine experimental methodologies and analytical procedures used throughout this work.

2.1 Chemicals and media

All chemicals used were purchased from Sigma Aldrich (Gillingham, UK) of analytical grade or higher. Bacteriological peptone and yeast extract were purchased from Oxoid (Basingstoke, UK). Reverse osmosis (RO) water and deionized water was used throughout for media preparation. A summary table of the media used is provided in 10.1.

2.2 Yeast strains

All *P. pastoris* transformants described in this work were provided by iQur Ltd. (London, UK).

Two phenotypes of *P. pastoris* were used throughout this work: Mut^{+} (X33) and Mut^{S} (KM71h).

P. pastoris Mut^{+} was used to express GFP-fused VLPs (GFP,e). *P. pastoris* Mut^{S} was used to express the following VLP constructs: k1,k1; LAH3,k1; HA2.3, $(M2e)_3$; and LAH3, $(M2e)_3$.

With the exception of the k1,k1 construct, amino acid sequences of tHBc-VLP constructs were proprietary at the time of this writing, and could therefore not be disclosed. Figure 2.1 shows the design of the k1,k1 construct.

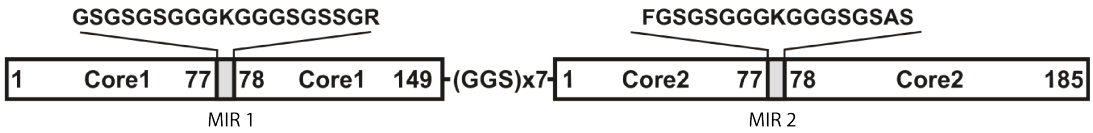

Figure 2.1 Design of the tHBc-k1,k1 construct. Amino acid sequences are depicted in the Major Insertion Regions (MIRs). Image was adopted from published work (Kazaks *et al.*, 2017).

2.2.1 Cell bank preparation

2.2.1.1 Master Cell Bank (MCB)

10µL frozen seed stock was cultured in 50mL Greiner (Greiner Bio-One, Kremsmünster, Austria), skirted centrifuge tubes in 9990µL BMGY media [13.8g/L yeast nitrogen base with ammonium sulphate and without amino acids, 12.6g/L glycerol, 100mL/L 1M potassium phosphate buffer pH 6.0, 20g/L bacteriological peptone, 10g/L yeast extract and 2mL/L biotin solution] for 20 hours at 30°C, 250 rpm in a Kuhner® (Glossop, UK) incubator. The optical density at 600nm (OD_{600}) was typically found to be 5 at this stage. 2mL of this culture was subsequently used to inoculate a 2L baffled polymer reusable Nalgene™ Erlenmeyer shake flask containing 250mL BMGY. Post-inoculation, OD_{600} values at this stage were typically around 0.2. The culture was harvested after 22-24 hours, typically reaching OD_{600} values between 60-80, indicating stationary phase. The harvested culture was diluted to an OD_{600} of 36 using BMGY and subsequently diluted to 25 using 100% sterile glycerol to achieve a final glycerol stock concentration at 30% (v/v). The suspension was vortexed briefly and distributed in 2mL aliquots. These aliquots were gradually frozen using a Nalgene® 5100-0001 cryo freezing container rack at -80°C for 24 hours.

2.2.1.2 Working cell bank (WCB)

A Master Cell Bank (MCB) vial was defrosted at room temperature for about 10 minutes. 2mL of each tube was transferred to a reusable Nalgene™ Erlenmeyer shake flask containing 250mL BMGY and grown as described in section 2.2.1.1. OD_{600} measurements were taken frequently to characterise the MCB and develop a growth profile (see Figure 10.1). To assess MCB quality, this procedure was carried out in duplicate. The harvested culture was processed as described in section 2.2.1.1 and will be referred to as the Working Cell Bank (WCB). This procedure was repeated without harvesting material to characterise the WCB (see Figure 10.2).

2.2.1.3 Agar plate quality control

As an additional quality control measure to ensure monocultures of transformants, agar plate-based culture inspection was performed. 50µL of frozen cell bank was grown on YPD agar plates [20g/L bacteriological peptone, 10g/L yeast extract and 20g/L agar] for one week at 28°C. Cultures were assumed to be of a monoculture when no colonies with different morphology (colours, hyphenation, shape and opacity) were observed (see Figure 10.3).

2.3 Fermentations

2.3.1 Microwell fermentations

As this thesis investigated the development of microwell culture protocols, a variety of methods were used rather than a single protocol. The commonalities of these procedures are described in this section.

Basal Salts Medium (BSM) [26.7mL/L 85%phosphoric acid, 0.93g/L calcium sulphate, 18.2g/L potassium sulphate, 14.9g/L magnesium sulphate heptahydrate, 4.13g/L potassium hydroxide, 40g/L glycerol, 20mL/L 28% ammonium hydroxide] was prepared without glycerol to serve as a variable induction medium. The volume of glycerol was substituted by reverse osmosis water. 4.35mL of PTM_1 trace salts [6.0g/L cupric sulphate pentahydrate, 80mg/L sodium iodide, 3.0g/L manganese sulphate hydrate, 0.2g/L sodium molybdate dihydrate, 20mg/L boric acid, 0.5g/L cobalt chloride, 20.0g/L zinc chloride, 65.0g/L ferrous sulphate heptahydrate, 0.2g/L biotin, 5.0mL/L sulfuric acid] was added to 1 litre of media. Media was adjusted to pH 5.0 and was filter sterilised using a 0.22µm filter.

An exponential inoculum culture (OD_{600}=10-30) was subjected to buffer exchange to achieve defined buffer conditions and inoculum concentration. The buffer exchange was performed by creating aliquots of equal volumes of seed culture into 15mL Greiner tubes (Greiner Bio-One, Kremsmünster, Austria) and centrifuging the aliquots at 3200g, 20°C for 10 minutes. The supernatant was discarded and

appropriate induction medium was added to each aliquot to achieve a target optical density followed by vortexing.

Variable volumes of the vortexed suspension were transferred to 24 well pre-sterilised deep square plates (Kinesis, St Neots, UK), which was sealed with sterile AeraSeal™ film.

2.3.1.1 Induction with Pectin Digest Medium

Some microwell fermentations involved induction using Pectin Digest Medium (PDM).

PDM[5] was prepared by autoclaving a suspension of 2.7mL 85% phosphoric acid, 8mL 50% ammonium hydroxide and 3.67g pectin from apple. The suspension was cooled to 25°C and the pH was adjusted to 5.5 using NaOH or 2M HCl. 30mg pectolyase Y-23 from *Aspergillus japonicas* (MP Biomedicals, Eschwege, Germany) was added to the suspension. The enzymatic reaction was left for 33 hours. BSM components and PTM1 trace salts were subsequently added in the appropriate amounts. The pH of the solution was adjusted to 5.0 as described earlier. The suspension was divided into two 50mL aliquots in 50mL centrifuge tubes. To remove solids, the suspension was centrifuged at 4°C 3200g for 10 minutes. The pellet was discarded. The supernatant was filter sterilised using a 0.22µm vacuum filter and stored at 4°C.

[5] 0.5% (v/v) MeOH molar equivalent d-Galacturonic Acid Methyl Ester (dGAME)

2.3.2 Shake flask cultures

A 2L reusable polymer baffled Nalgene™ shake flask containing 250mL BMGY was inoculated with 1.8mL working cell bank (2.2.1.2). This was incubated for 16-20 hours at 30°C and 250 rpm in a Kuhner® (Glossop, UK) incubator. When exponential phase characteristics were required, the culture was grown to OD_{600} values between 10-20. When stationary phase characteristics were required the culture was grown beyond and OD_{600} value of 60.

2.3.3 Mut^S 250 millilitre bioreactor fermentations

2.3.3.1 Pure methanol induction

A 250mL ambr® modular bioreactor system (Sartorius, Epsom, UK) was used for studying *P. pastoris* Mut^S fermentations. Each bioreactor unit (see Figure 2.2) had 5 reservoirs of which reservoir A contained 20mL 50% (v/v) glycerol solution supplemented with PTM1 trace salts; reservoir B 25mL 10% ammonium hydroxide solution; reservoir C contained 5mL sterile PPG 2000 antifoaming agent; 5mL 8.5% phosphoric acid was added to reservoir D; and reservoir E contained 25mL methanol with PTM1 trace salts. The bioreactors were heated to 30°C prior to calibrating Dissolved Oxygen Tension (DOT) and pH. After at least 3 hours, 2mL of sample was drawn from each of the bioreactors to perform offline pH measurements, which were required to complete pH calibration. Bioreactors were deemed ready for inoculation at DOT=100±5%, pH 4.75-5.0 and T=30±0.1°C. Bioreactors had a working volume of 100mL of which 7mL was provided by an

appropriately diluted and resuspended[6] inoculum culture to achieve a starting bioreactor culture optical density of 1.0 $OD_{600.}$ Throughout the fermentation DOT was maintained at 30% by (1) agitating the culture between 1323 and 3175 rpm, (2) maintaining an airflow rate of 0.5vvm and, (3) when required, additively blending pure oxygen with the sparged gas.

A 20% drop in Carbon Evolution Rate (CER) and spike in DOT, indicating depletion of carbon source, triggered a fed-batch glycerol feed. This was generally observed between 18-20 hours after bioreactor inoculation. This glycerol fed-batch phase was maintained for a fixed 4 hours at a constant flow rate between 11.9-24.4 millilitres per litre initial working volume per hour ($mL/L_i/h$). 20 minutes prior to the induction phase the temperature setpoint was varied between 18.8-31.3°C.

For the first two hours of the induction phase the methanol flow rate was kept constant at $1mL/L_i/h$. After this the feed rate was increased by 10% increments every 30 minutes until the target feed rate was reached. Note that the time of each final increment was shortened accordingly to achieve the precise final value of Q_{MeOH}. Also note that this feeding profile did not apply for a target flow rate of $0.5mL/L_i/h$. This was kept constant throughout the induction phase. After 40 hours, the culture was cooled to 12°C to minimise proteolytic activity. This was followed by harvesting (2.4.1).

[6] In BSM medium supplemented with PTM_1 trace salts.

2.3.3.2 Mixed feeding induction with glycerol

This method involved the same procedures as described in 2.3.3.1, however there were several differences. Instead of using methanol, mixed induction media was used comprised of a 60:40 ratio of 50% (*v/v*) glycerol and pure methanol respectively, plus 12 mL PTM_1 salts per litre of induction media. Bioreactors had a working volume of 83.3mL of which 6.16mL was provided by an inoculum culture (10-30 OD_{600}) to achieve a starting bioreactor culture optical density of 1.0 OD_{600}. Throughout the fermentation DOT was maintained at 30% by (1) agitating the culture between 1130 and 2826 rpm, (2) maintaining an airflow rate of 0.51vvm and, (3) when required, blending pure oxygen with the sparged gas, maintaining a constant gas flow rate.

The induction phase was followed immediately after the batch phase. Feeding of induction media was kept constant at 5mL/L_i/h, at a temperature of 30°C. Harvesting was performed after 48 hours of induction.

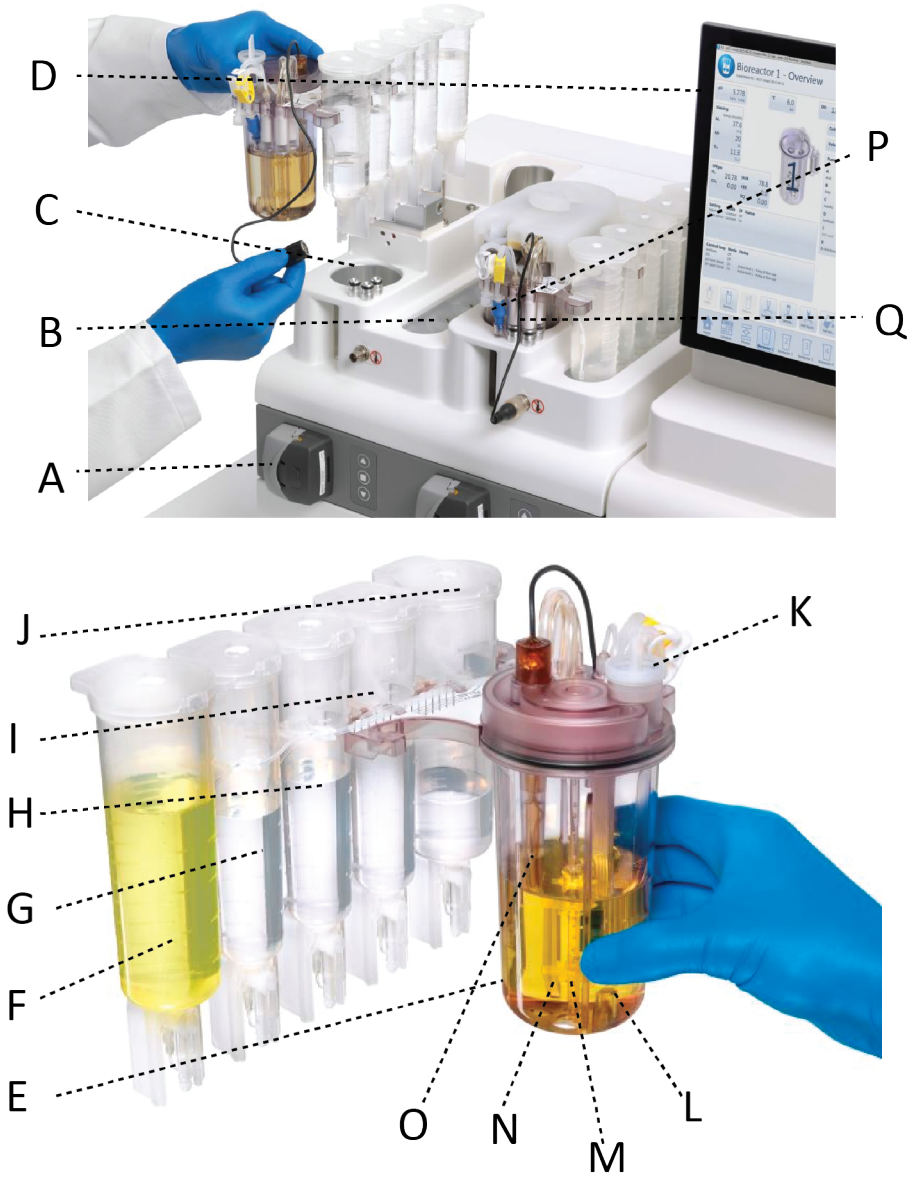

Figure 2.2 Overview of the 250mL ambr® modular bioreactor system. (A) Peristaltic pump for harvesting or additional feeding. (B) Reservoir slot. (C) Heating and cooling jacket. (D) Computer interface. (E) 250mL bioreactor with baffled walls. (F-J) Reservoirs A-E. (K) Removable stopper for inoculation procedures. (L) DO sensor. (M) Impeller with two Rushton turbines. (N) Gas sparger. (O) pH and temperature probe. (P) Sampling and addition lines. (Q) Gassing manifold.

2.3.4 Mut^S 7 litre bioreactor fermentations

A 7L BioFlo® New Brunswick (Eppendorf, Stevenage, UK) was used for initial fermentation scoping studies using the Mut^S pure methanol induction protocol (Invitrogen Corporation, 2002).

The initial working volume was set at 4L. The DOT setpoint was set at 30% and was controlled in a sequence cascade by agitating the impeller between 500 to 1200 rpm followed by oxygen gas blending in ratio mode at a constant volumetric gas flow rate of 1.5L/min. Fermentations were run in batch mode for 18-20 hours until a DOT spike was observed. This was followed by a 4 to 6-hour glycerol fed-batch phase at a constant glycerol feed (50% (v/v) glycerol solution supplemented with PTM1 trace salts) flow rate of 18.15 mL/L$_i$/h. Following this was the induction phase at 30°C. For the first two hours of the induction phase the methanol flow rate was kept constant at 1mL/L$_i$/h. After this the feed rate was increased by 10% increments every 30 minutes until a target feed of 3.0mL/L$_i$/h rate was reached. The induction phase was followed by harvesting as described in 2.4.2.

2.3.5 Mut^S 30 litre bioreactor fermentations

2.3.5.1 Pure methanol induction

A 30L BIOSTAT® Cplus (Sartorius, Epsom, UK) stainless steel bioreactor was used to study pilot scale cultivation and induction of Mut^S type *P. pastoris*. The bioreactor was run at 12L working volume of which 500mL was provided by an appropriately

diluted and resuspended[7] inoculum culture to achieve a starting bioreactor culture optical density of 1.0 OD_{600}. The DOT setpoint was set at 30% and was controlled in a sequence cascade by agitating the impeller between 468 to 1123 rpm followed by oxygen gas blending in ratio mode at a constant volumetric gas flow rate of 0.5vvm. pH range was maintained between 4.75-5.0 and pre-induction temperature at 30±0.1°C. A 20% drop in CER and spike in DOT, indicating depletion of carbon source, triggered a fed-batch glycerol feed. This was generally observed between 18-20 hours after bioreactor inoculation. This glycerol fed-batch phase was maintained for a fixed 4 hours at a constant glycerol feed flow rate of 18.15 $mL/L_i/h$. 20 minutes prior to the induction phase the temperature setpoint was adjusted to 25°C. The end of the glycerol fed-batch phase triggered the methanol induction phase. For the first two hours of the induction phase the methanol flow rate was kept constant at $1mL/L_i/h$. After this the feed rate was increased by 10% increments every 30 minutes until a target feed rate of $5.5mL/L_i/h$ was reached.

After 40 hours of induction the culture was cooled to 12°C to minimise proteolytic activity. This was followed by harvesting as described in 2.4.2.

2.3.5.2 Mixed feeding induction with glycerol

This method involved the same procedures as described in 2.3.5.1, however there were several differences. Instead of using methanol, mixed induction media was used comprised of a 60:40 ratio of 50% (*v/v*) glycerol and pure methanol respectively, plus 12 mL PTM_1 salts per litre of induction media. Bioreactors had a

[7] In BSM medium supplemented with PTM_1 trace salts.

working volume of 10L of which 500mL was provided by an inoculum culture (10-30 OD_{600}) to achieve a starting bioreactor culture optical density of 1.0 OD_{600}. Throughout the fermentation DOT was maintained at 30% by (1) agitating the culture between 400 and 1000 rpm, (2) maintaining an airflow rate of 0.51vvm and, (3) when required, blending pure oxygen with the sparged gas, maintaining a constant gas flow rate.

The induction phase was followed immediately after the batch phase. Feeding of induction media was kept constant at 5mL/L_i/h, at a temperature of 30°C. Harvesting was performed after 48 hours of induction.

2.3.6 Mut$^+$ 7 litre bioreactor fermentations

2.3.6.1 Pure methanol induction

A 7L BioFlo® New Brunswick (Eppendorf, Stevenage, UK) was used for initial fermentation scoping studies using the Mut$^+$ pure methanol induction protocol.

The initial working volume was set at 4L. The DOT setpoint was set at 30% and was controlled in a sequence cascade by agitating the impeller between 500 to 1200 rpm followed by oxygen gas blending in ratio mode at a constant volumetric gas flow rate of 1.5L/min. Fermentations were run in batch mode for 18-20 hours until a DOT spike was observed.

This was followed by a 4-hour glycerol fed-batch phase at a constant glycerol feed flow rate of 18.15 mL/L_i/h. The induction temperature was subsequently varied

between 20-30°C. The culture was induced with methanol, supplemented with 12mL/L PTM_1 trace salts. The initial flow rate was maintained a 3.6 mL/h per litre initial working volume (mL/L_i/h) for 2-4 hours depending on the stability of the DOT reading. The feed rate was subsequently increased to 7.3 mL/ L_i/h for another 2 hours and then increased to 10.9 mL/L_i/h for the remainder of the fermentation. The induction phase was followed by harvesting as described in 2.4.2.

2.3.6.2 Mixed feeding induction with sorbitol

This method involved the same procedures as described in 2.3.6.1, however there were several differences. The induction temperature was fixed at 20°C and a mixed 50:50 (c-mol:c-mol) ratio of sorbitol to methanol feed was introduced instead of pure methanol.

2.4 Preparative methods

2.4.1 Small-scale harvesting

This procedure applies to the harvesting of material derived from fermentation vessels of 250mL or less. Fermentation broth was harvested at 3200 g, 20 min and 4°C using a benchtop 5417R centrifuge (Eppendorf, Stevenage, UK) for samples of 50mL and a benchtop 5415R centrifuge (Eppendorf, Stevenage, UK) for samples of 2mL or less. Supernatant was discarded and the wet pellets were weighed and stored at -20°C.

2.4.2 Large-scale harvesting

This procedure applies to the harvesting of material derived from 7 and 30L bioreactors. Fermentation broth was harvested at 3000 g, 20 min and 4°C using a Beckman Avanti J-25 floor centrifuge (Beckman Coulter, High Wycombe, UK). Supernatant was discarded and the wet pellets were weighed and stored at -20°C.

2.4.3 Cell disruption with High Pressure Homogenisation (HPH)

Frozen cell paste was weighed and resuspended in MOPS lysis buffer [50mM 3-(N-morpholino)propane sulfonic acid (MOPS) , [1mM 4-(2-aminoethyl) benzene sulfonyl fluoride] (AEBSF) hydrochloride, 5U/mL benzonase (E8263-25KU), 5mM Dithiothreitol (DTT) in reverse osmosis water (R>18.2 MΩ/cm) titrated with 1M sodium hydroxide to achieve a of pH 7.5 (@10°C)] or TRIS lysis buffer [20mM Tris X , (1mM 4-(2-aminoethyl) benzene sulfonyl fluoride (AEBSF) hydrochloride, 5U/mL benzonase (E8263-25KU), 5mM Dithiothreitol (DTT) in reverse osmosis water (R>18.2 MΩ/cm) titrated with 1M sodium hydroxide to achieve a of pH 8.5 (@10°C)] to achieve wet cell weight (WCW) concentrations of 30 g/L unless stated otherwise. 40mL aliquots were prepared and subsequently disrupted using an APV Gaulin Lab40 High Pressure Homogeniser as stated between 300-1200 bar, for 1-5 passes. Homogenisation was performed at T<10°C with the aid of a glycol cooling loop.

To each homogenised sample a 1% volumetric addition of Triton X-100 stock solution was added to facilitate protein release. After a one-hour incubation period

at 4°C a 1% volumetric addition of EDTA stock solution was added to each aliquot. 1 mL from each aliquot was collected for clarification and subsequent analysis.

2.4.4 Cell disruption with sonication

Frozen cell paste was weighed and resuspended in TRIS lysis buffer (2.4.3). A SoniPrep 150 (Wolf Laboratories, Pocklington, UK) was used for cell disruption and was operated at an amplitude of 10 µm for three 6-second cycles with 10-second cool-down periods in between cycles.

After sonication, to each of the samples 10µL of 1/10 dilution triton-X100 solution was added to achieve a 0.1% v/v concentration. This was left to react for around 20 minutes prior to clarification steps (see 2.4.10).

2.4.5 Cell disruption with Adaptive Focused Acoustics (AFA)

An E210 AFA device (Covaris, Brighton, UK) was used to perform AFA-mediated cell disruption. Experiments were conducted with 1mL cell suspensions in MOPS lysis buffer, 0.1% Triton X-100, in 12 x 12mm milliTUBE vials (Covaris, Brighton, UK), each containing an integrated fiber to assist disruption. Alternatively, 6mL Chromacol vials (Thermo Fisher Scientific, Paisley, UK) were used. The tubes were secured in a 4 x 6 rack. The rack was placed into the water bath of the Covaris system containing degassed water at a temperature of 10± 1°C and a submerged acoustic transducer. Cell disruption was performed in power tracking mode, automatically by placing each tube in a predetermined sequence in the focal zone of the transducer. A 1% volumetric addition of EDTA stock solution was added after sonication.

2.4.6 Cell disruption with enzymatic lysis

MOPS Enzymatic Lysis buffer [4% (v/v) lyticase stock solution (2000U/mL lyticase from *Arthrobacter luteus* (*L2524*), 100mM potassium phosphate, 100mM sodium hydroxide, 50% (v/v) glycerol, pH 7.5 in RO water) and 1% (v/v) Triton-X100 stock solution in MOPS lysis buffer] was heated in a bath to 25°C. Frozen cell paste was resuspended in the warmed buffer to a variable target biomass value, followed by a 1 hour incubation at 25°C, as per manufacturer's instructions (Sigma-Aldrich, 2017).

2.4.7 Chemical lysis

Frozen cell paste was resuspended and diluted to an OD_{600} value of 5 using reverse osmosis water to a total volume of 1mL in 1.5mL micro-centrifuge tubes. The tubes were subsequently centrifuged at 2000 rpm for 10 minutes using a benchtop Eppendorf Centrifuge (5415R) to separate cells from the liquor. The supernatant was discarded and the pellets were resuspended in 100µL lysis buffer (20mM Tris, pH 8.0, 1mM EDTA, 5mM DTT, 0.75% Sodium Deoxycholate). The suspensions were incubated at room temperature for 20 minutes to react with lysis buffer as per findings of previous unpublished works (data not shown).

2.4.8 Cell disruption with YeastBuster™

Cell pellets were resuspended at room temperature in 5mL/g_{WCW} YeastBuster™ reagent (Merck Millipore, Watford, UK) and 50µL/g_{WCW} 100X THP solution. Benzonase was added as described previously 2.4.3. The suspensions were

incubated at room temperature for 20 minutes to react with YeastBuster™, as per manufacturer's instructions (Novagen, 2004).

2.4.9 Bead lysis

Bead lysis served as a small-scale cell disruption tool for iQur Ltd and was performed by IQur as outlined below.

Pellets were resuspended in 1.5ml MOPS lysis buffer and mechanically lysed through bead lysis (1g mixed bead (2 part 0.1mm beads (0.4g), 2 parts 0.5mm beads(0.4g) and 1 part 1.5-2mm beads (0.2g) ratio 2:2:1)) for 30 seconds on a vortex mixer followed by 30 seconds of cooling time on ice. This cycle was repeated 10 time per sample.

Lysate was separated from beads through manual extraction with a pipette. This lysate was sonicated using a SoniPrep 150 device, operated at 10000 amplitude microns for two 10 second cycles.

A 10% Triton X100 stock solution was added to achieve a final concentration of 0.1%.

2.4.10 Clarification

Lysates were centrifuged at 4°C and 15000 *g* for 30 minutes using a benchtop 5415R centrifuge (Eppendorf, Stevenage, UK) for samples of 2mL or less for sample volume of 2mL or lower. A benchtop 5417R centrifuge (Eppendorf, Stevenage, UK) was used

for sample volumes over 2mL. The supernatant of each sample was with a Millex 0.22µm 33mm PVDF syringe filter (Merck Millipore, Watford, UK). A Millex 0.45µm syringe filter was preceded by this if sample volumes were over 2mL. Filtrate was stored at 4-7°C for up to one week for immediate analysis; or stored long-term at -20°C.

2.4.11 Preparative Size Exclusion Chromatography (SEC)

Core-positive material within the size range of 20-200nm was purified through Size Exclusion Chromatography (SEC). The 200nm upper limit was assumed as the preceding step to this unit operation was consistently a 200nm depth-filtration step. The 20nm lower limit was defined as the assumed smallest size for a HBc VLP. Purification was achieved by passing 2mL of clarified lysate through a Sepharose CL4B column (XK16/20) fitted on an ÄKTA pure pump system (GE Healthcare Life Sciences; Hatfield, UK). The flow rate of the mobile phase (20mM Tris pH 8.4, 5mM EDTA, 1M Urea) was set to 1mL/min. Elution was collected between 10 and 14 minutes, as per recommendation from iQur Ltd., and is referred to as the 'pooled fraction' in Figure 2.3.

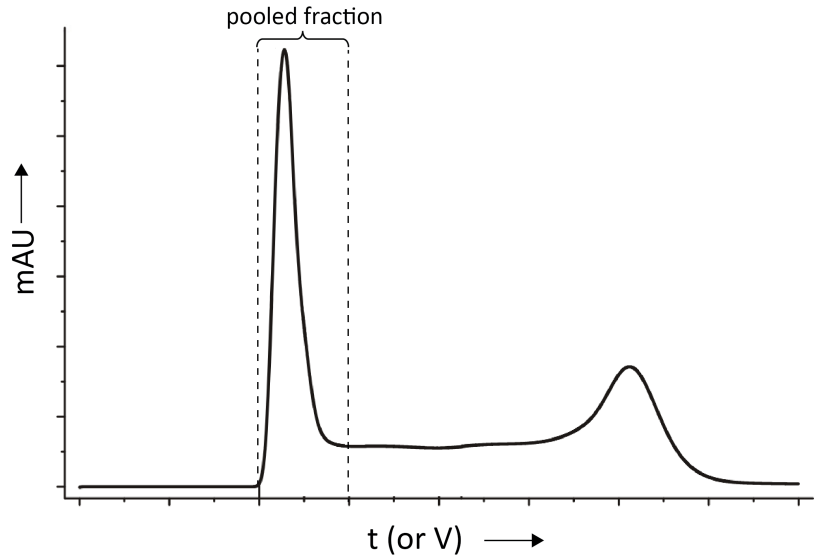

Figure 2.3 Typical size exclusion profile for tHBc-k1, k1. Adapted from other works (Kazaks *et al.*, 2017)**.**

2.5 Analytical methods

2.5.1 Biomass quantification using Optical Density (OD)

Optical density analysis of fermentation samples was performed with an Amersham Ultrospec 500 Pro spectrophotometer (GE Healthcare Life Sciences; Hatfield, UK), at an absorption wavelength of 600nm, using 1mL of diluted fermentation samples in acrylic cuvettes. These samples were diluted appropriately with deionised water until a reading between 0.1-0.5 was obtained. The optical density was determined by multiplying the reading with the corresponding sample dilution factor.

2.5.2 Biomass quantification using Wet Cell Weight (WCW)

When using small-scale bioreactors (V<250mL), wet cell weight (WCW) was determined as follows. In a pre-weighed 1.5mL Protein LoBind tube (Eppendorf,

Stevenage, UK), a 1.0mL sample was centrifuged at 3200 g, 10°C for 10 minutes using a benchtop Eppendorf Centrifuge (5415R). The supernatant was either discarded or collected for subsequent analysis. The weight of the pellet was determined by weighing the remaining pellet in the centrifuge tube and subtracting the weight of the tube.

When using large-scale bioreactors ($V_{W,i}$>10L), WCW was determined as follows. In a pre-weighed 50mL skirted Greiner tube (Greiner Bio-One, Kremsmünster, Austria) 50mL of sample was centrifuged at 3200 g, 10°C for 10 minutes using a benchtop Eppendorf Centrifuge (5417R). The supernatant was either discarded or collected for subsequent analysis. The weight of the pellet was determined by weighing the remaining pellet in the centrifuge tube and subtracting the weight of the tube.

2.5.3 Biomass quantification using Dry Cell Weight (DCW)

Dry cell weight (DCW) was only performed for small-scale fermentations ($V_{W,i}$<100mL). This was done by taking a pellet in a tube (see 2.5.2) and drying the tube in an oven at 100°C for 20-24 hours. The weight of the pellet was determined by weighing the dried pellet in the centrifuge tube and subtracting the weight of the tube.

2.5.4 Total protein analysis with Bicinchoninic Acid (BCA) assay

Triplicate two-fold dilution series of 25µL samples were prepared alongside a Bovine Serum Albumin (BSA) standard dilution series in a 96-well plate. 175µl of working reagent (50:1 ratio of BCA solution to a 4% $CuSO_4$ solution) was added to all

wells. The plate was wrapped in foil and incubated for 30 minutes at 37°C. Plate reading was performed at 562nm using a Safire II multi-plate reader (Tecan, Reading, UK). Based on a 1mg/mL protein upper detection limit, it was determined which sample dilution factor is appropriate for iterative BCA assays.

2.5.5 Total protein analysis with Nanodrop 1000

Total soluble protein analysis of unclarified and clarified lysates was performed using a Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific, Paisley, UK). Triplicate measurements were performed for 3µL sample volumes at an absorbance of 280nm per manufacturer's instructions. The extinction coefficient was set at 1.0 as a default general reference setting.

2.5.6 Nucleic acid analysis with Nanodrop 1000

Nucleic analysis of unclarified and clarified lysates was performed using a Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific, Paisley, UK). Triplicate measurements were performed for 3µL sample volumes at an absorbance of 260nm as per manufacturer's instructions.

2.5.7 Product quantification using Green Fluorescent Protein (GFP)

Green Fluorescent Protein (GFP) fused to Tandem Core protein, was measured in 100µL triplicate samples in a 96-well plate format. Plate reading was performed at an excitation wavelength of 483nm and an emission wavelength of 535nm, using a Safire II multi-plate reader (Tecan, Reading, UK).

2.5.8 Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

12 or 15-slot pre-cast NuPage™ 4-12% Bis-Tris Gels (Invitrogen, Carlsbad, USA) were placed in an XCell Surelock® Mini-Cell (Invitrogen, Carlsbad, USA). 200mL MOPS buffer (Invitrogen, Carlsbad, USA) was added to the inner chamber of the gel tank. 500µL NuPage™ antioxidant was added to the inner chamber.

30 µL of each sample was transferred to a 1.5mL micro-centrifuge tube with 10µL NuPage™ 4x LDS sample buffer and 4µL NuPage™ 10x reducing agent. All tubes were centrifuged for 30 seconds to ensure that all liquid resided at the bottom of the tubes. The tubes were vortexed and heated using a C1000 Touch™ Thermal Cycler (Bio-Rad, Watford, UK) at 95°C for 10 minutes to denature protein content.

3µL of pre-stained protein standard was added to the first lane of each gel. Denatured sample mixtures were applied to the gel slots at 10-15µL depending on the slot size. 600mL MOPS buffer was added to the outer chamber of the gel tank. Electrophoresis was performed at 200V for 50 minutes.

2.5.9 Western blotting

After SDS-PAGE, the gels were transferred to an XCell™ II Blot Module (Invitrogen, Carlsbad, USA) allowing for protein transfer to two respective nitrocellulose membranes (Invitrogen, Carlsbad, USA). Transfer buffer was prepared by mixing NuPage™ 20x Transfer Buffer (Invitrogen, Carlsbad, USA) with 749 mL ultrapure water, 200mL pure ethanol and 1 mL NuPage™ antioxidant. The transfer was run at

10V for 3 hours. For membrane development, the immunostaining protocol (section 2.5.11) was followed.

2.5.10 Dot blotting

Samples were diluted in MOPS lysis buffer (2.4.3) to achieve total protein concentrations of 1mg/mL at a total volume of 50µL in 0.2mL Protein LoBind™ PCR tubes (Eppendorf, Stevenage, UK)

Reference standard (Recombinant Hepatitis B Core Antigen, Abcam AB49013, 1.0mg/mL) was diluted in MOPS 0.1% TX100 buffer to achieve 100µL stock solutions. These stock solutions were serial diluted two-fold to final volumes of 50µL to generate reference curves. 5µL 10X NuPage™ sample reducing agent was added to all aliquots. These were subsequently heated at 95°C for 10 minutes using C1000 Touch™ Thermal Cycler and subsequently cooled for at least 15 minutes at 12°C.

For each assay, a 0.45µm 11.5 x 8.5 cm nitrocellulose membranes was placed on a piece of filter paper (12 x 9cm) resting on a table. Using a 12 x 8 roster to serve as a grid, samples were applied on the membrane in 2µL volumes. Each membrane was dried for at least 10 minutes at room temperature post sample application.

For membrane development, the immunostaining protocol (section 2.5.11) was followed.

Note that, unless stated otherwise, all samples were analysed with triplicate measurements to determine average values and standard deviations.

2.5.11 Immunostaining

Membranes were blocked in 5% skimmed milk power PBS-T (0.05% Tween20) solution for 1 hour at room temperature or overnight at 4-7°C. Hepatitis B virus core antigen-specific, mouse monoclonal antibody (Abcam; AB8639; Cambridge, UK) was applied in a 1:1000 solution in 2.5% Skimmed Milk Power PBS-T and incubated for 45 minutes at room temperature and washed three times for 5 minutes each wash with PBS-T. Secondary antibody (Abcam; A4416-1ML; Cambridge, UK) was applied in a 1:2000 dilution in PBS-T followed by a 30-minute incubation at room temperature. Membranes were finally washed three times for 5 minutes with PBS-T and once for 5 minutes with PBS to rinse away residual detergent. 10mL of Bio-Rad's Clarity Western ECL substrate was used to develop the membranes. Detection was performed using automated exposure setting on an Amersham Imager 600 (GE Healthcare Life Sciences; Hatfield, UK). Dot blot densitometry was done using ImageQuant software followed by quantification with four-parameter logistic (4PL) fitting.

2.5.12 Particle Size Distribution analysis with static light scattering

Static light scattering analysis was performed using a Mastersizer 3000 (Malvern Instruments, Malvern, UK). Samples were applied to the Hydro SV Dispersion unit at

1000rpm until laser obscuration was in the 5-10% range. 5 measurements of each sample were taken to subsequently create an average PSD curve.

2.5.13 Methanol analysis using High Performance Liquid Chromatography (HPLC)

Methanol concentrations in clarified supernatant samples were determined by HPLC using a Dionex UltiMate3000 HPLC pump and fitted with an Aminex HPX-87H column. Columns were run at 0.6 mL/min for 30 minutes using 0.1% (v/v) trifluoroacetic acid (TFA) as mobile phase, a column temperature of 60 °C, an injection volume of 10 µL and monitored with an ERC Refractomax 520 refractive index detector. Quantitative analysis was performed by integration of peak areas using the external standard method (Ward *et al.*, 2015).

2.5.14 Neutral sugar analysis using High Performance Liquid Chromatography (HPLC)

Clarified supernatant samples were analysed using a Reagent-Free™ Ion Chromatography System (ICS) (Thermo Fisher Scientific, Paisley, UK) fitted with an AminoPac PA10 (2 × 250 mm) anion exchange column, with AminoPac PA10 guard column (2 × 50 mm), an eluent generator with a KOH 500 cartridge, and an electrochemical detector (gold electrode). An injection volume of 10 µL was used and a 7.5 mM KOH solution was used as the mobile phase with a flow rate of 0.25 mL/min for 18 min at 30 °C. Quantitative analysis was performed by integration of peak areas using the external standard method (Ward *et al.*, 2017).

2.5.15 Uronic acid analysis using High Performance Liquid Chromatography (HPLC)

Clarified supernatant samples were analysed as described in section 2.5.14 however, instead using a mobile phase of 5% (v/v) 1 M aqueous sodium acetate and 95% (v/v) deionised water at 0.25 mL min^{-1} for 5 min at 30 °C (Ward *et al.*, 2017).

2.5.16 Transmission Electron Microscopy (TEM)

TEM was performed by qualified operators of the Latvian Biomedical Research and Study Centre.

2.5.17 Phase-contrast and fluorescence microscopy

Unclarified lysate samples were diluted to [X]=5-10g_{wcw}/L using PBS and stained using a 1:1000 volumetric addition BODIPY® in DMSO. After an incubation period of 15 minutes, 5µL aliquots of stained suspension were used to prepare microscopy slides. Microscopy analysis was done in bright light microscopy and fluorescent imaging modes using an Eclipse microscope (Nikon, Kingston upon Thames, UK).

2.5.18 Statistical analysis and Design of Experiments (DoE)

All statistical analysis and model generation was done using JMP®Pro12.0.1 (SAS Institute Incorporation; Cary NC, USA).

Screening models were generated with a minimum resolution of 5 to estimate all possible two-factor interactions. The relative contribution of an individual factor was defined as the total of the sum of squares (Type III) of each factor-associated term as a percentage of the total of the sum of squares of all terms in the screening models.

Central Composite Designs (CCDs) with on-face axial points (α=1) and two center points were used to generate quadratic RSMs. Unless stated otherwise, only statistically significant model terms ($p<0.05$) were included in each model and were selected using stepwise regression.

Chapter 3 Initial process characterisation of novel Tandem Core VLP constructs

3.1 Introduction

The aims of bioprocess development are to create processes that produce the highest possible yield of product, at the highest purity, at maximum throughput, with the greatest level of reproducibility at a variety of scales, and, preferably at the lowest possible cost.

As discussed in the chapter 1, when aiming to fulfil the above criteria, there exist multiple process options to produce a single biopharmaceutical compound. Although the scope of this research has been limited to using, specifically, *P. pastoris* as an expression system to produce Tandem Core Virus-Like Particles (VLPs), a multitude of process options still need to be addressed.

Firstly, one must consider the architecture of the plasmid to be tested. As the Tandem Core VLP is still novel, testing for appropriate assembly and expression of the VLP will be required. In this case, rather than choosing an immunogenic fusion protein associated with influenza, reporter proteins, such as Green Fluorescent Protein (GFP) or Luciferase, may provide better insight and quick verification of tandem-core protein presence. In this work, this group of variables will be referred to as construct-associated variables.

Secondly, the expression system needs to be chosen appropriately. Variables in this area are the specific type of host cell that would be suitable, typically considering host cell mechanisms that will facilitate appropriate formation of product; and what mutations have been introduced to achieve desired metabolic effects. In this work, this group of variables will be referred to as host cell-associated variables.

Thirdly, there are several upstream process options available. The first aspect to consider is which platform is most suitable. In upstream microbial bioprocessing this typically involves a variety of available modes of fermentation: either continuous mode, batch mode, fed-batch mode, or a combination of these modes of fermentation. A subsequent aspect to consider is how the chosen platform is operated. Operating variables typically include operating temperature, agitation rate, aeration or sparge rate, operating pH, Dissolved Oxygen Tension (DOT), induction initiation time and harvest time. Media composition also needs consideration and can be controlled through media formulation, culture feeding and culture dilution. For fed-batch systems, feed flow rates need to be considered, whereas with continuous systems, culture dilution rates need consideration. This group of options will be referred to as upstream process-associated factors.

Fourthly, there is a multitude of options available for primary recovery. Primarily, and indeed in the case of this research, the product of interest needs to be liberated from the cells. This can be achieved with various disruption methods: autolysis, chemical lysis, enzymatic lysis, bead-facilitated lysis, sonication, adaptive focused acoustics and homogenisation. Each of these methods has trade-offs in

terms of throughput, scalability and effectiveness. Notably, sonication and enzymatic lysis are not used at pilot scale because no large-scale apparatus exists for the former and operational costs are too high for the latter.

Subsequently, a suitable method must be chosen to separate cell debris from the soluble product of interest. This can be achieved through a variety of combinations of various modes of centrifugation and filtration. Again, these modes of operation have various operating parameters associated with them that can considerably affect yield, purity, throughput, scalability, reproducibility and cost. This group of process factors will be referred to as primary recovery-associated factors.

Finally, the recovered product can be purified in numerous ways. For the purification of heterologous protein, this typically involves removal of unwanted protein species and nucleotides. This can be achieved through numerous precipitation methods, gradient centrifugation, and many chromatographic methods, including varieties of size-exclusion chromatography, hydrophobic interaction chromatography and ion-exchange chromatography. This group of process factors will be referred to as downstream process-associated factors.

Due to the sheer magnitude of options available, process-design limitations were imposed. Construct-associated factors were limited to the investigation of upstream process effects by several fusion protein inserts. This is because cell-engineering specifically falls outside the scope of this research. For the same reason, host cell-associated factors were limited to investigating relevant differences in the use of

Mut^S-type (slow methanol metabolising) and Mut^+-type (fast methanol metabolising) *P. pastoris*, specifically and respectively the KM71h and X33 strains. Upstream process-associated factors were limited to the use of a combination of batch mode and fed-batch mode fermentations, as described by Invitrogen (Invitrogen Corporation, 2002). The chosen primary recovery-associated factors, excluding extraction methods, were restricted to bench-top centrifugation and depth filtration methods. Downstream process-associated factors were constrained to salt precipitation methodology and to Size Exclusion Chromatography. These two unit operations were previously characterised for the purification of monomeric HBc protein (Strods *et al.*, 2015).

3.1.1 Hypotheses and outline

Due to the relative novelty of this work, a significant portion of previously discussed process-associated factors could not be addressed through literature alone.

This chapter therefore aims to identify critical process parameters and to generate observations through addressing relatively simple initial hypotheses. These observations can be used to (1) formulate more in-depth hypotheses and (2) establish a base case experimental platform which, in turn, can be used to address these matured hypotheses.

The order of the hypotheses and the associated research reported in this chapter will generally follow the same sequence as unit operations in a bioprocess: addressing upstream-associated factors, by investigating fermentations; followed

by addressing primary recovery-associated factors by investigating various modes of cell disruption; followed by addressing downstream-associated factors.

3.2 *Pichia pastoris Mut*$^{+}$ expressing Tandem Core protein with GFP insert

Previous works have shown that Tandem Core Hepatitis B Virus-Like Particles can be produced in both bacteria and plants (Peyret *et al.*, 2015); while other VLPs, such as those based on Human Papilloma Virus (HPV), have been produced in *P. pastoris* using bioreactors (Jiang *et al.*, 2011). Initial work was undertaken to establish fermentation protocols for the correct expression and formation of Tandem Core in *P. pastoris*, and to assess formation using a GFP reporter for analysis. This reporter protein greatly increases the ease of detecting the presence of Tandem Core protein throughout a bioprocess in comparison to flu epitope inserts, which would require more elaborate assays such as ELISA's or immuno-densitometry, and a good supply of reagents.

This section will report on how *P. pastoris* X33 (Mut^{+}) was used to express Tandem Core VLPs with a single GFP insert on one of the Major Insertion Regions (MIRs). Using the Mut^{+} phenotype should provide strong expression of heterologous protein when compared to its Mut^{S} variant. Note that a key assumption made here is that higher expression of Tandem Core protein will result in higher VLP yields.

3.2.1 Bench-top fermentation of *Pichia pastoris* Mut$^+$

The main aims of this experiment were to investigate a *P. pastoris* X33 fermentation to identify critical process parameters and to study Tandem Core HBc-GFP protein (tHBc-GFP,e) expression and stability over time, using Western blotting.

All stages of fermentation were operated as described in section 2.3.6.1, with the following exceptions: as shown in Figure 3.1A, due to an extended period of poor temperature control, a higher than desired temperature (43°C) resulted in unfavourable conditions for biomass accumulation and, therefore, most likely prolonged the batch phase of the fermentation, although no product was induced at this stage. Additionally, the glycerol fed batch phase involved the use of a 50% w/w feed as opposed to a 50% volumetric feed, resulting in a slightly more dilute feed of non-inducing carbon source. Furthermore, the fed batch phase was 4.43 hours in length and was followed by a brief carbon starvation period to ensure the absence of residual glycerol prior to induction. Finally, an unintended pH drop was observed during the induction phase after approximately 40 hours of fermentation. This was caused by depletion of base reservoir however, this did not seem to have a significant impact on biomass concentrations.

OD_{600} measurements were taken as described in section 2.5.1 and can be seen in Figure 3.1B. In addition to biomass monitoring, using optical density, multiple samples were collected from the point of induction throughout the remainder of the fermentation.

Collected fermentation samples were diluted to 5.0 OD_{600} using PBS for the purpose of obtaining data relating to specific expression, rather than volumetric expression. Following centrifugation, the pellets were chemically lysed (section 2.4.7). Following lysis, the pellets were vortexed and 50µL samples were taken, representing the total protein expression fraction. This was followed by centrifugation so a that soluble protein fraction could be studied. To visualise specific protein expression over time, Western blot analysis was used, as described in sections 2.5.8 and 2.5.9. Immunostaining was performed as described in 2.5.11. Sigmafast™ 3,3-DAB tablets were used for detection (see Figure 3.2).

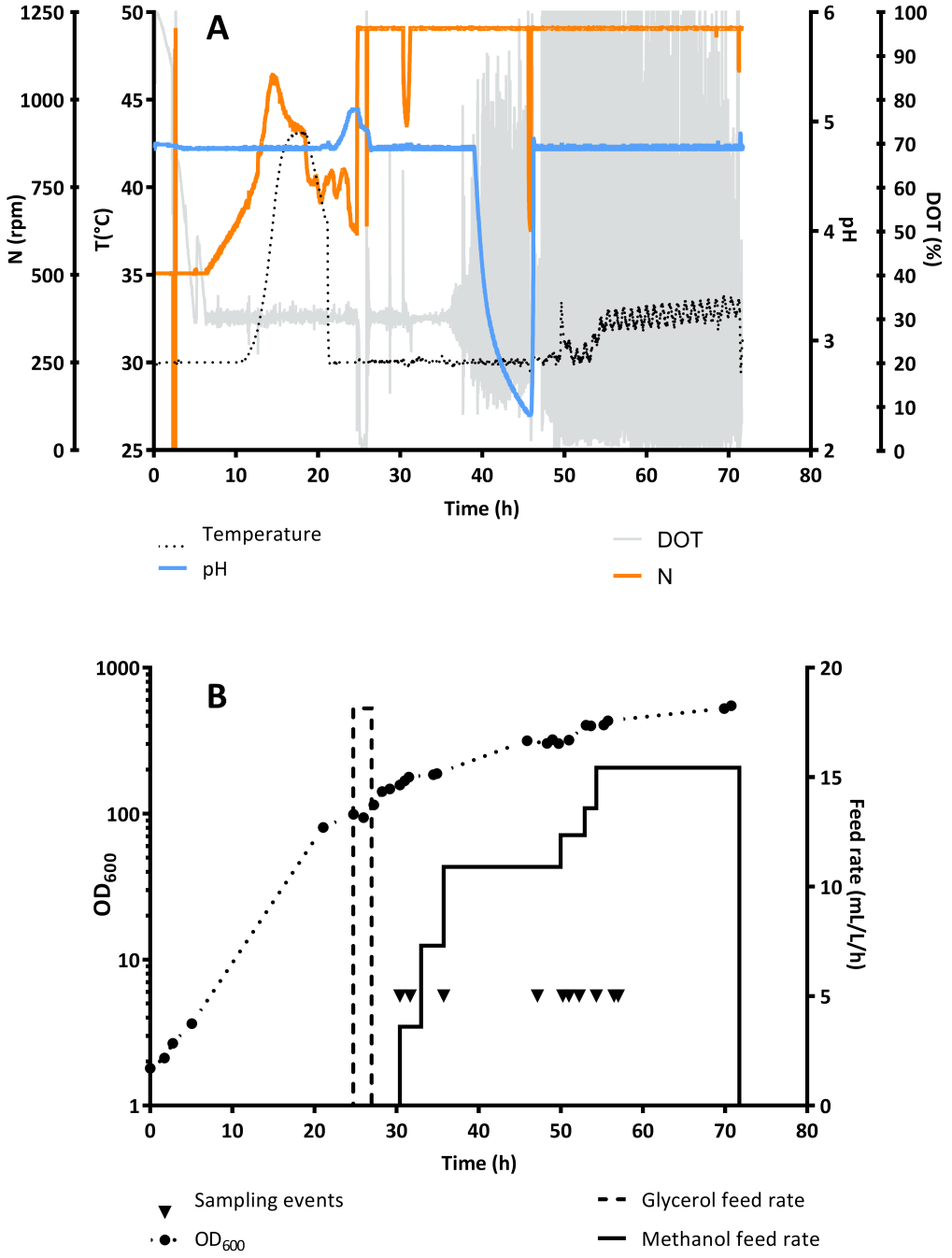

Figure 3.1 Fermentation data of *P. pastoris* Mut$^+$ expressing tHBc-GFP,e: (A) Online data including agitation rate, temperature, pH and dissolved oxygen tension level. (B) Offline data including OD_{600}, sampling events and carbon source feed rates.

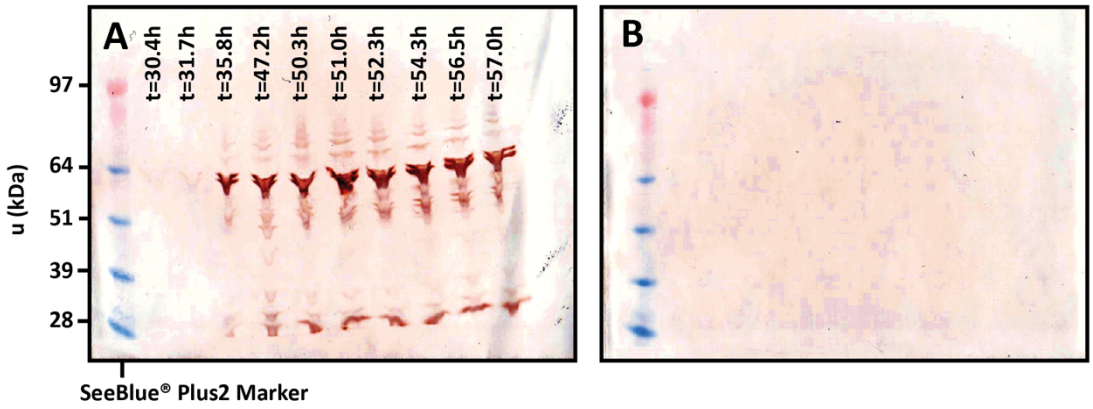

Figure 3.2 Western blots of reduced gels of collected and processed fermentation samples detected with hepatitis B core 10e11 antibody and stained with Sigmafast™ DAB: (A) Total tHBc-GFP,e expression over time. (B) Total soluble tHBc-GFP,e expression over time.

The fermentation showed to produce relatively large amounts of biomass, reaching 380g/L WCW at the point of harvest. The Western blot results also show that induction was successful. The full-length monomer is expected at a molecular weight of around 64kDa and, while this represents the main band, it is clear to see that some truncations of the monomer are visible, possibly indicating proteolytic activity. Furthermore, the higher bands may also indicate dimer formations, though this is difficult to establish with the existing resolution. However, unfortunately, most of the product was not recoverable in soluble form at any time of the fermentation. As the Western blots consistently showed degradation profiles, as opposed to a single product band, and as no appropriate reference hepatitis B core reference standard was used, it was not immediately evident at this point whether tHBc-GFP,e was expressed as intended.

3.2.2 Investigating alternative modes of induction and primary recovery

The previous section demonstrated that cultivation and induction of *P. pastoris* X33 expressing tHBc-GFP,e was possible. However, the majority of expressed heterologous protein detected was insoluble and therefore it was unclear whether VLP assembly had occurred. Furthermore, if VLP assembly had occurred, VLPs were not able to be detected due to potential association with insoluble cell debris, possibly as a result of the method of extraction employed.

Various sources indicate that changing operating temperature and implementing a sorbitol-methanol mixed feeding strategies can have beneficial effects on process yields and product quality (Li *et al.*, 2001; Xie *et al.*, 2005; Celik, Calik and Oliver, 2009; Zhong *et al.*, 2014). Figure 3.2 showed variations between band intensities and profiles over time, particularly when comparing early induction time points (t=30.4) to later induction time points. This indicates that induction time is a significant process parameter. Furthermore, other sources suggest that the mode of extraction does indeed have a significant effect on the recovery of VLPs (Kee, Pujar and Titchener-Hooker, 2008; Gurramkonda *et al.*, 2013).

Therefore, based on literature, in an attempt to improve soluble tHBc-GFP,e yields, this section will seek to test the following hypotheses:

1. Induction temperature has a significant effect on tHBc-GFP,e expression and quality, based on degradation profiles.

2. The implementation of mixed feed induction has a significant effect on tHBc-GFP,e expression and quality.
3. Duration of induction time has a significant effect on protein expression and quality.
4. tHBc-GFP,e yields are affected by different extraction methods.

To test the first hypothesis, a fermentation was operated as described in section 3.2.1, however, induced at a lower temperature of 20°C. Fermentation data is shown in Figure 10.4 in the appendix.

Looking at Figure 3.3, and comparing induction with pure methanol at 20°C with pure methanol induction at 30°C (as described in section 3.2.1), it appears that higher total soluble tHBc-GFP,e expression occurs at higher temperatures. Higher temperatures are producing more GFP, however it is unclear if the GFP is part of the VLP monomer or if it has been cleaved.

The Western blot in Figure 3.4B indicates that a higher amount of full-length tHBc-GFP,e monomer (band 2, ~64kDa) has become insoluble at a low temperature induction process, while higher induction temperatures seem to result in higher concentration of degradation product (bands 3 and 4). Therefore, the first hypothesis of this section seems to be proven: induction temperature has a effect on tHBc-GFP,e expression and quality. In this case, lower temperatures are shifting more of the product into the insoluble fraction.

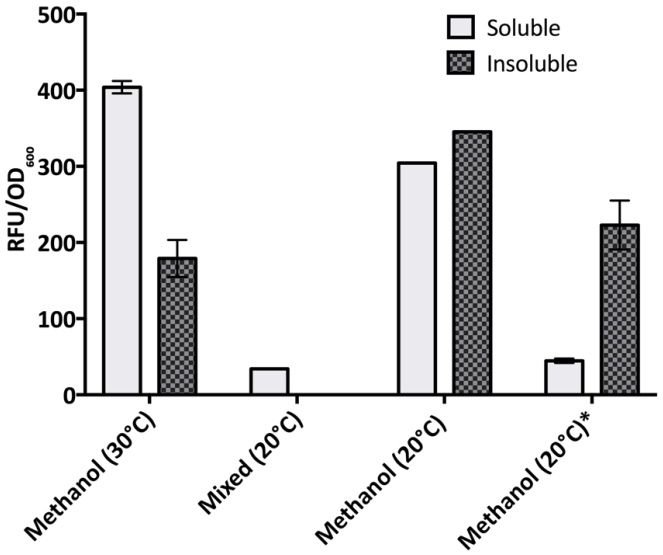

Figure 3.3 Specific tHBc-GFP,e expression (RFU/OD_{600}) of harvested material from three fermentations, each with different modes of induction. 'Methanol' refers to an induction process with only methanol in the feed; 'Mixed' refers to an induction process with a mixture of sorbitol and methanol in the feed. Extractions were performed with homogenisation or with sonication (*). Soluble and insoluble heterologous protein content was quantified with a GFP assay as described in section 2.5.7.

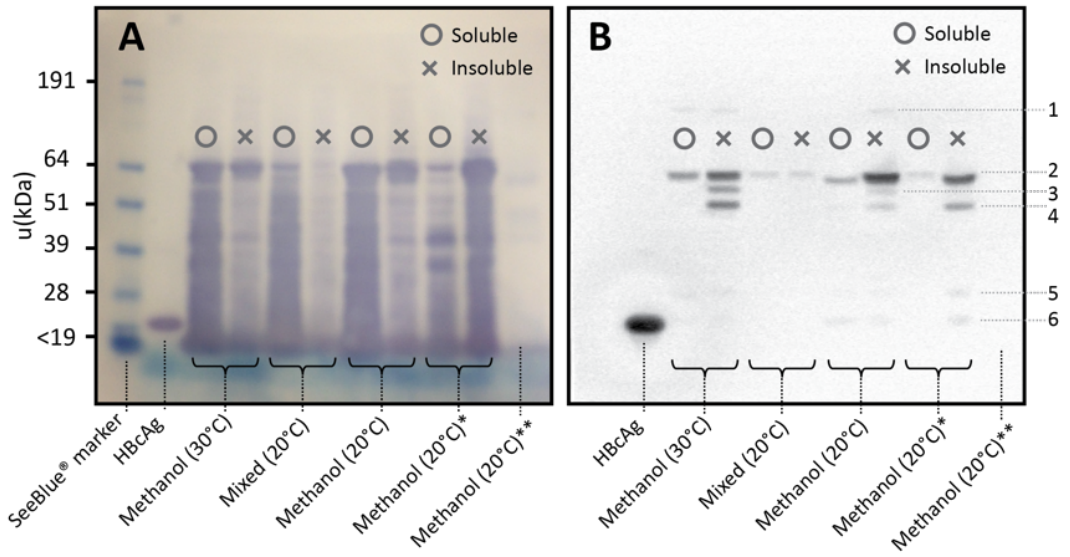

Figure 3.4 Fermentation sample analysis. Harvested material, each with different modes of induction. Extractions were performed with homogenisation or with sonication (*). Protein presence in clarified culture media was also studied (**). (A) Reversible membrane stain of reduced SDS-PAGE gel showing soluble and insoluble host cell protein expression profiles. (B) Western blot showing soluble and insoluble expression profiles of tHBc-GFP,e. Six bands were identified, assuming u,GFP=27kDa (Hink *et al.*, 2000)**, u,HBcAG=21kDA hetero-Tandem Core: (6): u≈21kDa, (5) u≈ 26kDa, (4) u≈51kDa, (3) u≈53 kDa,(2) u≈64 kDa and (1) 64<u<191kDa.**

The second hypothesis was addressed by running a fermentation as described earlier in this section, however, inducing with a 1:1 carbon mole ratio of sorbitol to methanol mixture, at an induction temperature of 20°C. Fermentation data is shown in Figure 10.5 in the appendix.

Comparing this fermentation with the previously discussed process where a 20°C pure methanol induction strategy was used, looking at Figure 3.3 and Figure 3.4, it appears the addition of sorbitol as a co-substrate does not obviously benefit the expression of soluble tHBc-GFP,e. In fact, expression of heterologous protein, per cell, is decreased through the addition of a co-substrate, possibly due to competition between the activation of the AOX1 gene and cell growth and maintenance. This could have resulted in lower specific heterologous expression. Nonetheless, of what little protein expression is monitored, a higher proportion of expressed protein appears to be soluble. Thus, while mixed feed strategies can indeed have an impact on expression of heterologous protein, the titres generated did not result in a beneficial improvement.

The third hypothesis was investigated by analysing protein expression of a low-temperature fermentation over time. As seen in Figure 3.5 and Figure 3.6, representing specific protein expression over time in the previously discussed fermentation with a 20°C pure methanol induction, time does appear to be a significant factor for overall expression of tHBc-GFP,e. However, note that the amount of soluble heterologous protein, measured in RFUs, is consistent over time: 8.27$\pm$1.61% relative to the total amount of heterologous protein expressed. This

data would suggest that time is not strictly a significant contributing factor to product solubility. Yet, looking at the corresponding Western blot (Figure 3.6B), it seems that as induction time progresses, insoluble degradation products (bands 2,3 and 4) accumulate. This conflicting data between RFUs from GFP and Western blot bands indicating HBc positive material seems to suggest that the GFP insert is cleaved from the Tandem Core scaffold. However, this is not a conclusive finding. In future work, this theory could be tested by employing GFP-specific antibodies in an orthogonal Western blot assay. Nonetheless, the conclusion of this experiment was twofold: (1) the hypothesis that induction time has a significant effect on protein expression and quality was proven; and (2) the GFP insert does not provide suitable characteristics for bioprocess characterisation.

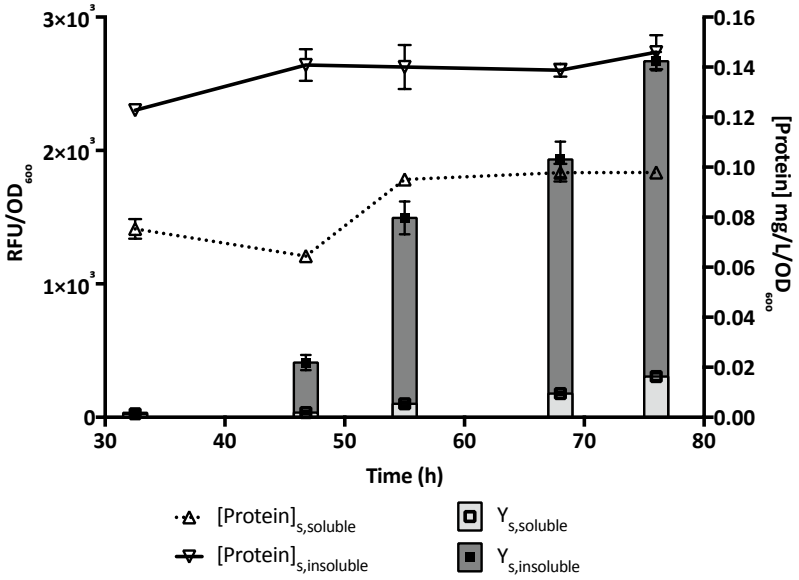

Figure 3.5 Time course analysis of protein expression of 20°C pure methanol induction. GFP-positive protein expression is displayed as Relative Fluorescence Units (RFUs) per unit of biomass (OD_{600}). Total protein is expressed in mass per unit of biomass.

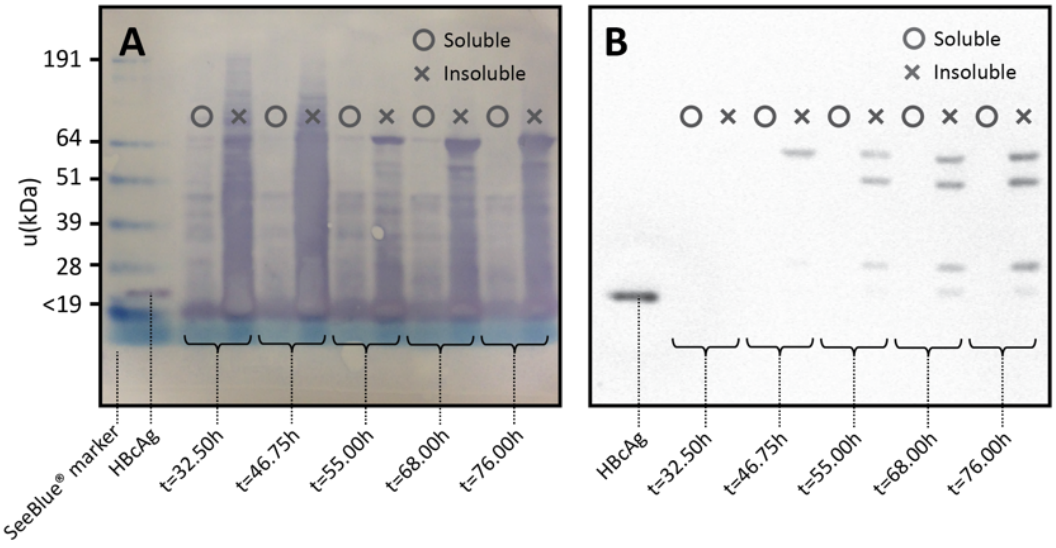

Figure 3.6 Protein analysis of fermentation time course. Extractions were performed with sonication. (A) Reversible membrane stain of reduced SDS-PAGE gel showing soluble and insoluble host cell protein expression profiles. (B) Western blot showing soluble and insoluble expression profiles of tHBc-GFP,e.

The fourth hypothesis was addressed by comparing protein extraction with homogenisation (1200 bar, 3 passes) (section 2.4.3), to protein extraction with sonication (section 2.4.4). Figure 3.3 and Figure 3.4 indicate that using sonication as an extraction method, results in lower overall yields and lower tHBc-GFP,e solubility compared to High Pressure Homogenisation.

In addition to assessing the above hypotheses, the following observations were made: The banding profiles on the Western blot on Figure 3.4B seem to indicate that truncation of heterologous protein contributes to insolubility. Six species, or breakdown products, of expressed heterologous tHBc-GFP,e protein fractions can be identified, numbered as 1 to 6. As seen on Figure 3.4, only species 1, 2 and 6 are found to be soluble. This suggests either certain truncations are inherently insoluble or so strongly hydrophobic that they associate with cell debris post

homogenisation; or that these specific truncated species affect the assembly, from a kinetic perspective, into soluble VLPs. Following electron microscopy analysis, no proof was found of VLP assembly in any of the soluble fractions of the lysed harvest pellets shown in Figure 3.4 (data not shown).

Due to inconsistencies in generating the analysed materials, it should be stressed that the findings in this section are strictly indicative, rather than conclusive. These irregularities include: temporary loss of control of operating parameters such as DOT, pH and temperature; different induction feed rates and different harvesting times. Harvesting times, ranging from 70 to 92 hours post-inoculation, were based on consistent target OD_{600} values of 500 rather than consistent harvesting time, which, after testing the third hypothesis of this section, seems to have proven to be an incomplete approach as time was shown to affect heterologous expression. Also note that volumetric productivity was not assessed as all final harvest densities were consistent approximately 500 ± 31 OD_{600}.

3.3 *Pichia pastoris* Mut^S expressing Tandem Core protein with lysine inserts

Section 3.2 described how attempts were made to characterise the production of VLPs consisting of Tandem Core protein units with linked GFP reporter as an insert on the Major Insertion Region. These units were expressed in a *P. pastoris* Mut$^+$ phenotype. Results show that, despite investigating multiple process parameters, most of the expressed material remained insoluble and was often degraded. This indicated that VLP formation had been problematic. In addition, results indicated

that the GFP-fused protein did not provide suitable characteristics for bioprocess characterisation. Indeed, no evidence was found of VLP production by the studied fermentations via electron microscopy. This could be because the GFP insert could have directly affected the assembly kinetics of VLPs, through mechanisms of steric hindrance and/or hydrophobic interactions. Secondly, the rate of heterologous protein expression, manipulated through varying modes of induction, was shown to have an effect on tHBc-GFP,e expression and solubility.

Work by Zlotnik *et al* indicates that VLP self-assembly requires regulation, as uncontrolled assembly of free units of recombinant protein can result in 'kinetic traps' and thus incomplete formation of capsids (Zlotnick and Mukhopadhyay, 2011). Specifically, at high induction rates, there will be multiple monomers that will seek to come together to form an intermediate VLP, but may never create a fully formed capsid. Instead, Zlotnik suggests that by having fewer monomers, the proportion of incomplete intermediate VLP will be reduced and result in more fully formed capsids. Therefore, the Mut^S phenotype of *P. pastoris* was utilised to assess this theory.

The *P. pastoris* Mut^S phenotype has a significantly lower methanol metabolism and thus also a lower rate of expressing heterologous protein (Cereghino and Cregg, 2000). As seen in section 3.2.2, where a mixed feed induction was compared to a pure methanol induction fermentation, lower specific expression seemed to result in relatively higher ratios of soluble to insoluble recombinant protein.

Following these observations, the next experiments tested whether Tandem Core protein assembly is hindered by the size of the insert placed in the MIR. It was noted that GFP appeared to cleave off the Tandem Core monomer, therefore this was replaced with a lysine residue in each MIR. Secondly, if VLP self-assembly requires regulation and can be controlled through lower protein expression rates, then *P. pastoris* Mut^S should be able to produce higher levels of appropriately assembled VLP than its higher metabolic counterpart, Mut^+.

3.3.1 Fermentation of *Pichia pastoris* Mut^S

Fermentations of the Mut^S phenotype of *P. pastoris* (KM71h strain) were conducted using a lysine residue insert at both MIRs on the Tandem Core VLP. This protein will be referred to as the k1,k1 construct, or tHBc-k1,k1. Relevant fermentation data is shown in Figure 10.6 in the appendix.

Prior to induction, the fermentation processes were similar between Mut^S and Mut^+. However, the induction phase yielded very different fermentation performance, mostly relating to methanol consumption. As expected, due to lower AOX expression, the Mut^S fermentation required a longer adaptation period to methanol and was not capable of utilising this substrate at the same rates as Mut^+ cultures. This is reflected by lower agitation rates in Mut^S cultures during the induction phase. To accommodate lower methanol consumption rates, preventing methanol toxicity, an induction method specific to the Mut^S genotype was used as described in section 2.3.4 (Invitrogen Corporation, 2002). However, it was expected that lower methanol consumption would also result in lower specific expression of

Tandem Core protein. This lower, and arguably more controlled, expression was postulated to result in higher VLP yields.

3.3.2 Initial DSP of VLP material

Harvested material was processed as described by Figure 3.7, based on protocols already established in literature for hepatitis B core VLP purification. (Freivalds *et al.*, 2011; Strods *et al.*, 2015)

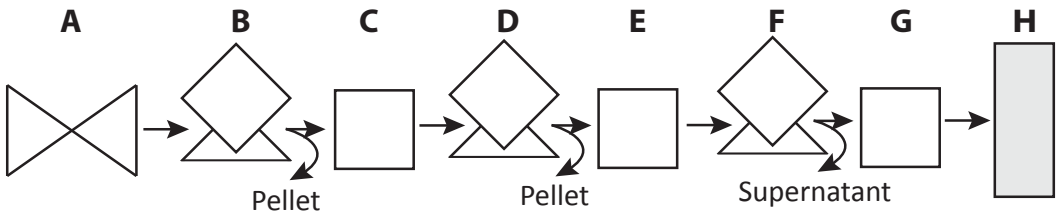

Figure 3.7 Summary of unit operations involved in purifying and characterising expressed material. (A) Homogenisation at 1200 bar, 3 passes (section 2.4.3); (B) centrifugation at 15000 g, 30 min, 4°C; (C) heat treatment in a water bath at 55°C, 30 minutes; (D) centrifugation at 15000 g, 30 min, 22°C; (E) ammonium hydroxide precipitation; (F) centrifugation at 15000 g, 30 min, 4°C; (G) suspend pellet in PBS; (H) preparative size exclusion column (section 2.4.11).

The use of a heat treatment step allowed for the isolation of heat stable particles in the supernatant. The implementation of ammonium sulphate treatment allowed for the precipitation of VLPs, which are relatively hydrophobic in comparison to other soluble protein species. Applying a final size-exclusion chromatography step, as described in section 2.4.11, allowed for the isolation of particles in the desired size range.

Figure 3.8B shows how this sequence of purification steps was used to successfully isolate Tandem Core protein with no, or negligible amounts of, truncated core-positive species. The full-length monomer is expected at molecular weight of around 42.5kDa. Bands 1 and 2 have been assumed to be full-length tHBc-k1,k1 proteins, however; band 1 is most likely a dimeric form of this monomer. Previous work has shown that this intermediate is required for VLP assembly and would be less susceptible to chemical reduction (Holmes *et al.*, 2015). Hence, this dimer could be present due to incomplete chemical reduction in preparation of the assay. Figure 3.9 shows how these same unit operations facilitate the removal of host cell protein and nucleic acids (Figure 3.9). Finally, Figure 3.10 proves that the expressed tHBc-k1,k1 monomers are assembling into Virus-Like Particles. As shown by the profiles associated with clarification in Figure 3.8B, it would appear that most of the expressed tHBc is soluble. A slight majority of this soluble material is heat-stable and, in turn, all of this material precipitated after the addition of ammonium sulphate. This last step eliminated a significant portion of host cell protein and nucleic acids, as shown in Figure 3.9, therefore proving it to be a highly effective purification step. Although the SEC step was shown to reduce the amount of HCP, it is not exactly clear why a large amount of core-positive material was lost during to this process. This could possibly be due to the detection of relatively small higher-order core-positive structures, or of individual units of unassembled Tandem Core protein that are removed during the SEC and pre-SEC filtration steps.

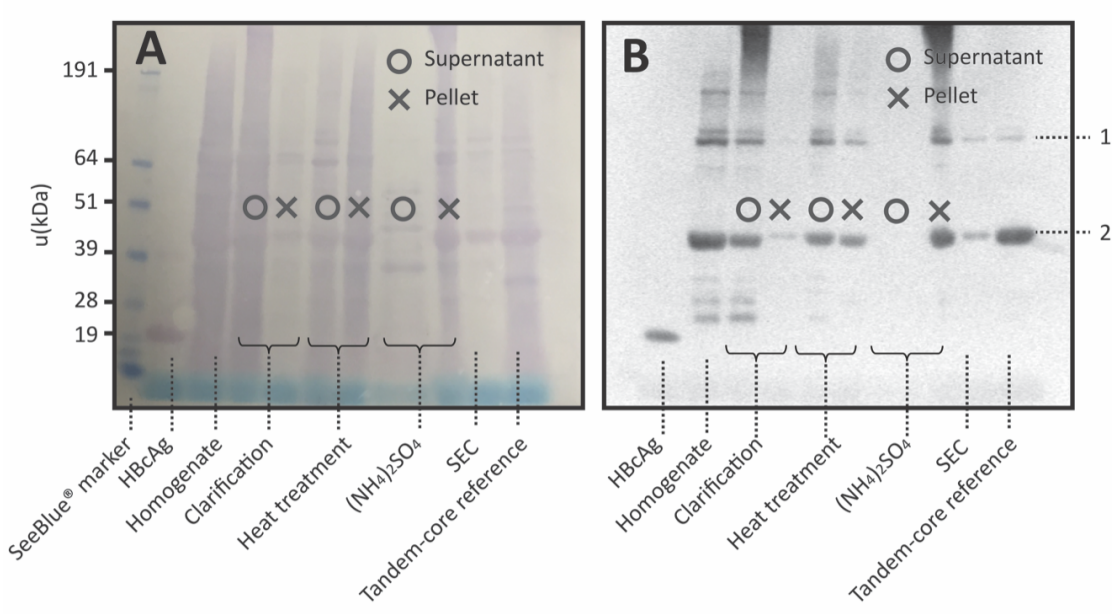

Figure 3.8 Qualitative protein data of purification sequence as described by Figure 3.7. (A) Membrane stain of a reduced SDS-PAGE gel showing host cell protein expression profiles. (B) Western blot showing expression profiles of tHBc-k1,k1. Bands 1 and 2 depict soluble, non-truncated Tandem Core protein species.

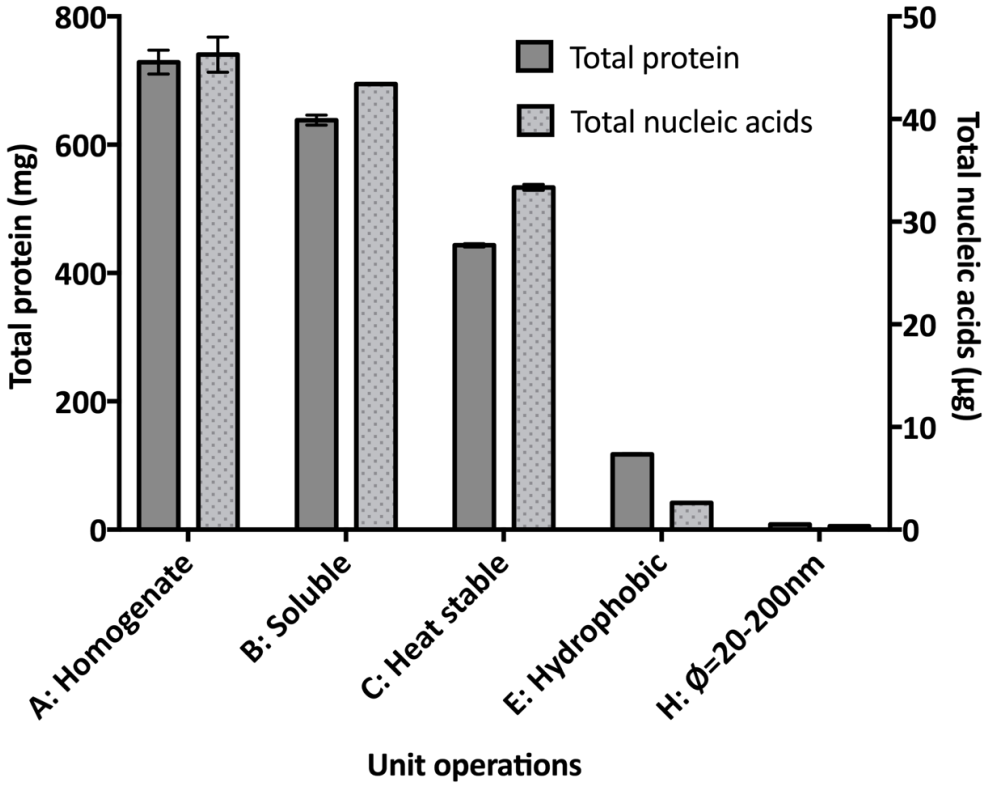

Figure 3.9 Mass balance of total host cell protein and total nucleic acids in the purification sequence described by Figure 3.7.

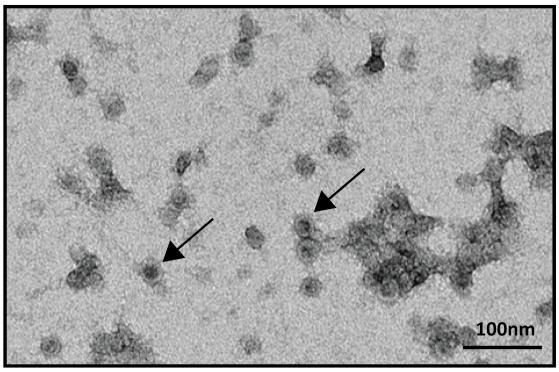

Figure 3.10 Electron microscopy image of k1,k1 Virus-Like Particles in a resuspended pellet after ammonium sulphate precipitation. VLPs are indicated by the arrows.

3.3.3 Investigating process robustness

The previous section demonstrated that tHBc-k1,k1 was successfully expressed, in a soluble form, and was assembling into Virus-Like Particles.

However, not addressed were: (1) the effect of time on the expression of soluble tHBc-k1,k1 and; (2) how robust this process is from a bioprocessing perspective.

To address this, firstly, a replicate fermentation run was attempted, the original fermentation being depicted previously in Figure 10.6. The data for this replicate run is shown in Figure 3.11. Samples were taken throughout the fermentation, which were then processed as described in Figure 3.7 until unit operation 'G', prior to Size Exclusion Chromatography, together with samples taken throughout the original fermentation. Expression of tHBc-k1,k1 in both the original and replicate fermentation were compared and were found to differ in band intensities (Figure 3.12).

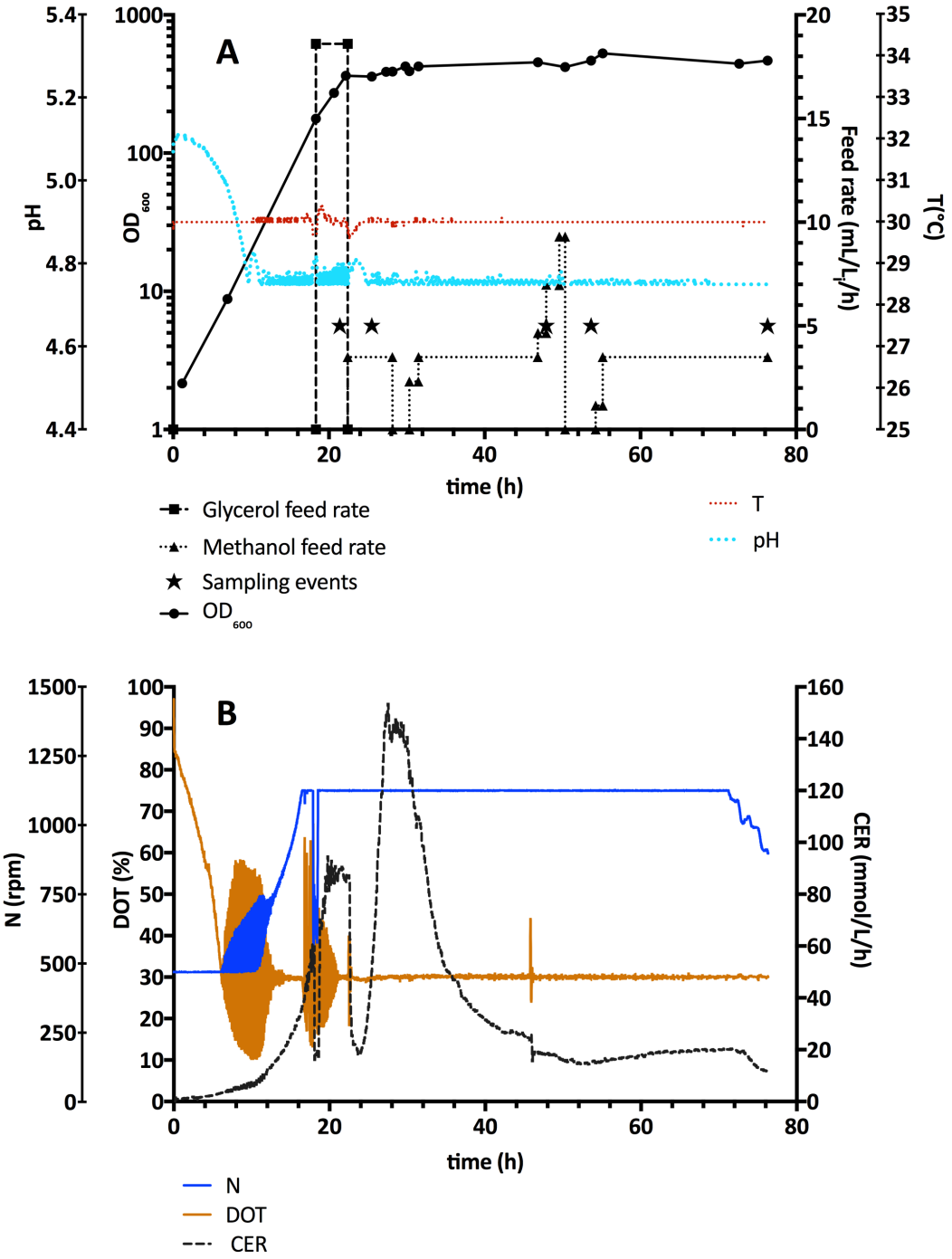

Figure 3.11 Data for replicate fermentation attempt of *P. pastoris* Mut^S expressing tHBc-k1,k1: (A) Offline data including OD_{600}, sampling events, carbon source feed rates, temperature and pH (B) Online data including agitation rate, dissolved oxygen tension level and carbon evolution rate.

Unfortunately, the fermentations could not be operated identically due to differences in discrete pump increments. However, this presented an interesting opportunity to assess upstream process robustness.

The main cause of differences in metabolic performance (CER and biomass) is most likely due to a 9% increase in volumetric feeding rates of methanol relative to the original fermentation. This explains the overfeeding-related CER profile seen in the attempted replicate fermentation, resulting from methanol accumulation. This profile is distinctive by an initial rapid increase in CER, between 21 and 24 hours of fermentation, followed by a rapid decrease in CER after 25 hours of fermentation.

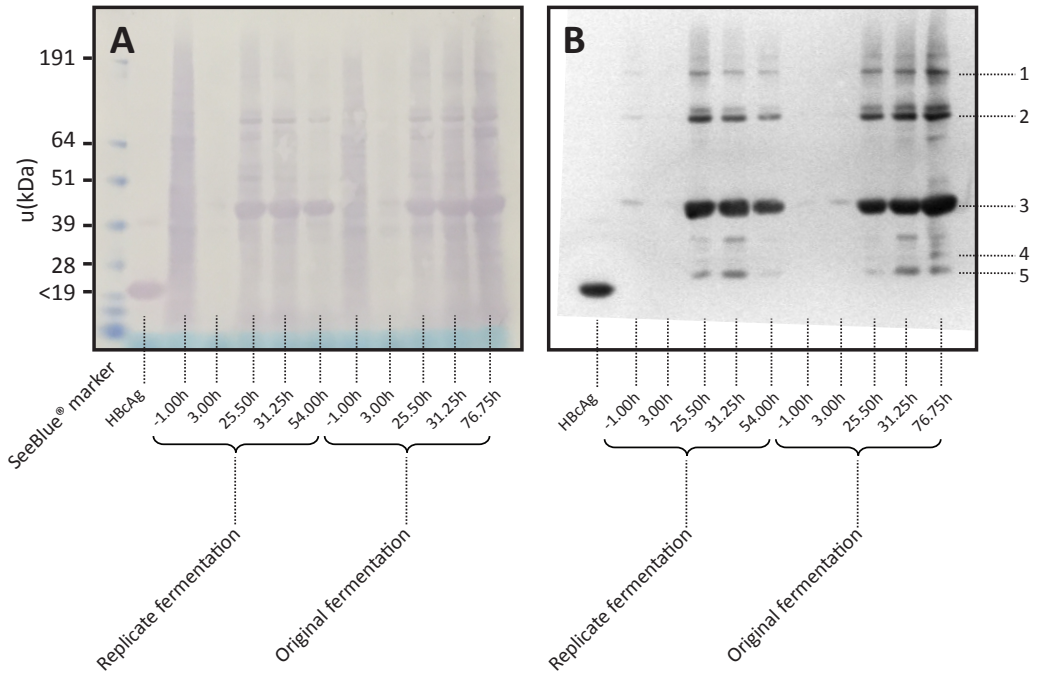

Figure 3.12 Qualitative protein data of time course analysis. Time represents the time post-induction. Extractions were performed as shown in Figure 3.7 up until step G. (A) Reduced SDS-PAGE gel showing soluble host cell protein expression profiles. (B) Western blot showing soluble expression profiles of tHBc-k1, k1.

Overall, it seems that a relatively small deviation (9%) in pump flow rates can have a significant effect on the fermentation process described. This process therefore

does not appear to be very robust as there is a major issue with upstream process reproducibility. Additionally, although ammonium sulphate precipitation facilitated significant removal of various impurities, it was found that the process could not be applied consistently. In particular, it was often difficult to isolate a pellet with centrifugation following precipitation (Figure 3.13). Therefore, future investigation should be carried out on more robust platforms with tighter control.

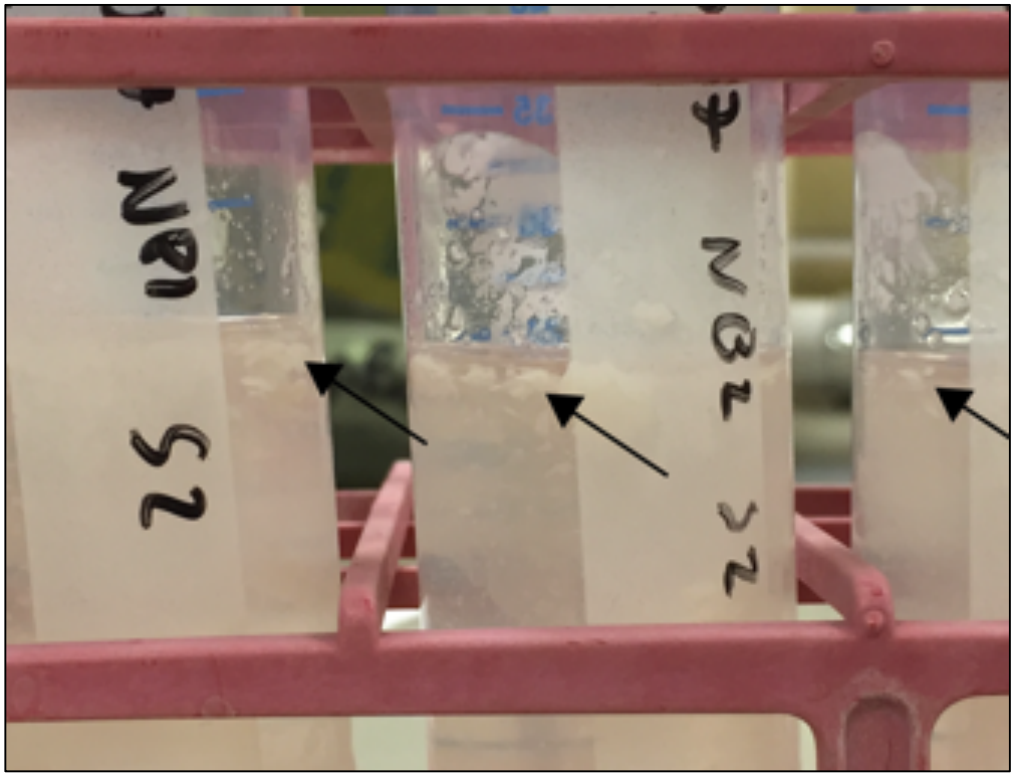

Figure 3.13 Close-up photo of 50mL tubes containing non-pelleted precipitate (highlighted by arrows) following ammonium sulphate treatment and centrifugation.

3.4 Summary

This chapter has shown that the production of recombinant tandem hepatitis B core (tHBc) VLPs is possible in suspension culture.

Initially, a Mut^{+} phenotype of *P. pastoris* was used to express Tandem Core with a single GFP insert in one of the MIRs. No evidence was found to suggest that the expressed material was assembling into VLPs. This could possibly be explained by the large number of product truncations detected in analyses. These truncations could have hindered the assembly of VLPs.

It was found that induction temperature, the addition of a sorbitol co-substrate during induction; and that induction time have a significant effect on Tandem Core expression and quality, with low quality indicated by high levels of truncations.

It was hypothesised that lowering expression rates through alteration of host cell-associated factors, and introducing smaller inserts, would result in more favourable conditions for the expression and assembly of soluble VLPs. This hypothesis was tested using a *P. pastoris* Mut^{S} strain expressing Tandem Core protein with two, single lysine inserts. Results showed that, using this strategy, soluble and heat-stable VLPs were indeed produced in the appropriate size range.

It must be stressed that certain results in this chapter, particularly those relating to upstream process-associated factors, are strictly indicative rather than conclusive. This is owing to a lack of consistency and reproducibility in experiments. In addition to this, minute changes to operating parameters such as feeding flow rates, caused drastic changes to product quality attributes.

To gain conclusive and quantifiable results, these findings need to be addressed in a robust, reproducible and, preferably, high-throughput fashion. This could be

achieved through the use of miniaturised fermentation platforms, more streamlined extraction processes, and combined with experimental design methodology.

The time and cost benefits of these miniaturised fermentation platforms can be gained only by employing complementary techniques, facilitating high-throughput at small sample volumes. As indicated earlier in this chapter, the method of extraction has a significant effect on the yield of soluble heterologous material and, since the investigated miniaturised extraction platforms investigated in this chapter yielded poor soluble yields, this is something that needs to be addressed in the subsequent chapter.

This chapter has established the formation of a VLP using the k1,k1 construct and a Mut^{S} phenotype strain. The following chapters will seek to conduct process development on these Tandem Core VLPs using a high-throughput methodology for upstream processing and primary recovery. As sample analysis has proven to be a limiting factor, lysis methodologies will be first investigated, which will facilitate sample analysis for high-throughput fermentation.

Chapter 4 Characterisation, miniaturisation and optimisation of primary recovery methods

4.1 Introduction

Chapter 3 demonstrated the significance of primary recovery-associated factors in the recovery of soluble Tandem Core material. Firstly, it was shown that the chosen extraction method has a significant effect on the yield of soluble material. Secondly, it was found that small-scale extraction methods did not allow for the recovery of relatively high soluble yields of core-positive protein.

The latter finding presented a significant challenge as the benefits of miniaturised fermentation platforms can only be gained by employing complementary techniques facilitating high-throughput at small sample volumes.

This chapter will therefore describe research aimed at structurally addressing the effects of primary recovery methods, with an emphasis on cell disruption, on the yield of soluble Tandem Core material. In doing so, the chapter will discuss how cell disruption processes could be characterised, optimised and miniaturised such that they could be employed in a platform approach when coupled with miniature fermentation, or situations of limited sample volumes. Note that all biomass (*P. pastoris* Mut^S expressing k1,k1 VLPs) was generated from a single batch as described in section 2.3.5.1. The majority of methodologies and results presented in this chapter have previously been published (Bláha *et al.*, 2017).

4.2 Methods of cell disruption

Cell disruption is essential for almost all intra-cellular products. Most studies have defined total cell disruption (*R*) as a performance criterion of cell disruption processes, often measured as total soluble protein content. However, counter-intuitively, maximising cell disruption is not always the main objective of cell disruption, as product can associate with cell debris, such that is becomes difficult to remove using centrifugation. This chapter will define cell disruption performance with five different performance criteria: (1) Total cell disruption, measured as the concentration of total soluble protein, *R (mg/mL),* (2) specific cell disruption, defined as the amount of protein released per unit of suspended biomass, R_s *(mg/g_{WCW})*[8], (3) total product recovery expressed as the concentration of soluble Tandem Core Hepatitis B core protein *[tHBc] (µg/mL),* (4) specific product recovery, $tHBc_s$ *(µg/g_{WCW})*[9] and (5) product purity, expressed as a ratio of total recover product relative to the amount of total release soluble protein, *P (%)*[10]. By understanding the response models of multiple performance criteria, the range applications of cell disruption platforms can be broadened as well as create a performance-based scaling methodology.

[8] $R_S = R/[X]$
with [X] being the concentration of biomass (g/L)
[9] $tHBc_s = [tHBc]/[X]$
[10] $P = [tHBc]/R$

4.2.1 High Pressure Homogenisation

The most widely-used method of cell disruption in microbial bioprocessing is High Pressure Homogenisation (HPH) (Middelberg, 1995; Kleinig and Middelberg, 1997; Klimek-Ochab *et al.*, 2011). This method involves passing a cell suspension at high pressure, and often low temperature to mitigate heat effects, through an adjustable valve with a restricted cavity as shown in Figure 4.1.

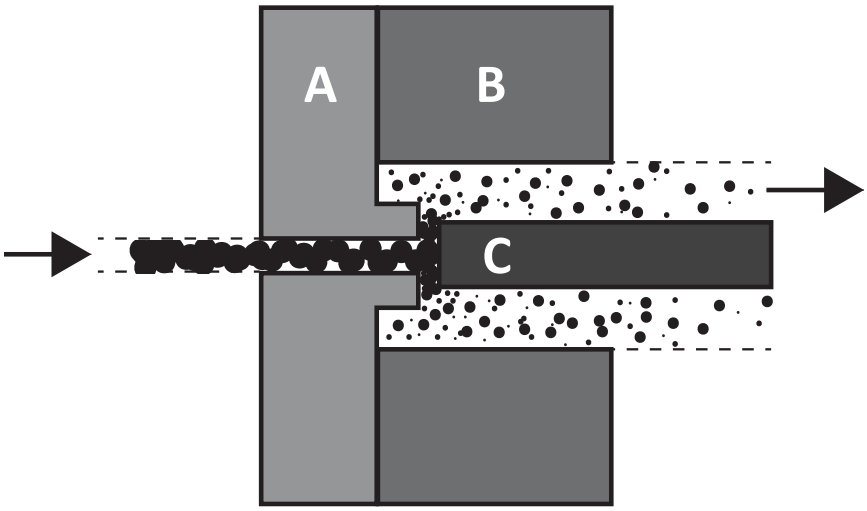

Figure 4.1 Schematic representation of a valve of a high-pressure homogeniser. (A) Valve seat, (B) impact ring and (C) valve rod.

Through high-velocity impact, cavitation, fluid shear and decompression, the cells are disrupted causing their contents to be released. This process can be repeated for several passes, increasing the levels of cell disruption and debris generated (Middelberg, 2000; Doran, 2013). Overall cell disruption can be described with the following equation (Follows *et al.*, 1971):

$$ln\left(\frac{R_{max}}{R_{max}-R}\right) = kNp^{\alpha} \quad \textbf{Equation 4-1}$$

Where, R_{max} is the maximum amount of releasable protein (mg/mL), *R* the observed amount of released protein (mg/mL), *N* the discrete number of passes through the valve, *k* a temperature dependent rate constant and is specific to the organism being disrupted; *p* the operating pressure (bar) and α is a measure of a microbe's resistance to disruption.

In addition, efficiency of homogenisation is known to decrease at high biomass suspension concentrations (Kleinig *et al.*, 1995). Therefore, for a given organism harvested at fixed conditions the following relation can be observed for cell disruption:

$$R = f(N, p, [X]) \quad \textbf{Equation 4-2}$$

Where *[X]* is the concentration of biomass (g_{WCW}/L) in suspension.

4.2.2 Miniaturised cell disruption platforms

However, High Pressure Homogenisation is only suitable for processing large volumes with the smallest representative size being 40mL using an APV Gaulin Lab40 homogeniser (Jenning, Lippacher and Gohla, 2002). Scale-down bioprocess sequences often start with the use of microtiter plates or miniaturised bioreactors (Kensy, Engelbrecht and Büchs, 2009b; Slinsby and Dewar, 2015). HPH therefore remains an incompatible platform for scale-down bioprocesses, as sample volumes are often in the microliter range.

This chapter will therefore investigate scale-down cell disruption platforms.

4.2.3 Aim and objectives

The aim of this chapter is to develop a miniaturised cell disruption platform intended for rapid upstream process characterisation. The specific objectives are to:

- Investigate the effects of lysis buffer components. This will be necessary as these factors could contribute to overall cell disruption, product release, product stability and, once optimised, will be invariable between large-scale and small-scale (mechanical) disruption platforms.
- Characterise high-pressure homogenisation by investigating a number of responses. This will be necessary to establish process performance targets for small-scale methods.
- Investigate the performance and implementation of several small-scale cell disruption technologies. This will allow for the selection of the most appropriate platform.
- Screen for significant process factors for the chosen small-scale cell disruption platform. This process will eliminate the need to investigate potentially insignificant process factors during optimisation studies, allowing for a more economical experimental approach.
- Characterise and optimise small-scale cell disruption process parameters. This will maximise process the output of key quality attributes and, in so

doing, could create response range-overlap with large-scale cell disruption performance.

- Investigate scalability and mimicry

4.3 Results

4.3.1 Buffer Studies

Separating buffer-associated variables from disruption process operation variables will allow for a more structured and formalised experimental approach. This chapter will aim to investigate the effects of various non-ionic detergents and salt concentrations on VLP recovery.

Buffer studies were conducted using fixed homogenisation conditions (3 passes at 1200 bar), as described in section 2.4.3.

All detergents investigated in this chapter are non-ionic, implying that the hydrophilic region of the molecule is uncharged. Typically considered a mild class of detergents, these affect lipid interactions without significantly affecting protein-protein interactions. This class of detergents is predominantly used to isolate membrane-associated proteins by forming micelles with the lipid components of the membrane (Shi *et al.*, 2005; Johnson, 2013). The three detergents investigated are Triton-X 100, Polysorbate 80 and Tween 20 as these have been studied for the recovery of other types of VLPs (Shi *et al.*, 2005). These non-ionic detergents differ in cloud point, critical micelle concentration, density and molecular weight.

After determining the most suitable type of detergent, a full-factorial screening was performed on the concentration of salt, the concentration of the chosen type of detergent and the concentration of biomass in suspension (also referred to as the concentration of solids).

Figure 4.2 shows the experimental procedures involved in studying the effects of various buffer conditions on a variety of key quality attributes. The studied attributes were (1) total cell disruption, measured as total protein concentration (section 2.5.5) after unit operation 'E'; (2) total Tandem Core material recovery, estimated using western blotting and densitometry (sections 2.5.9 and 2.5.11), measured after units 'E' and 'G'; (3) particle-size distribution of clarified lysate, measured after unit operation 'E'. Note that product analysis on western blotting is done after unit operation 'E' and 'G'. The rationale of this was that aggregated Tandem Core material would not pass through a 0.22µm filter and would therefore result in a decreased signal on Western blot.

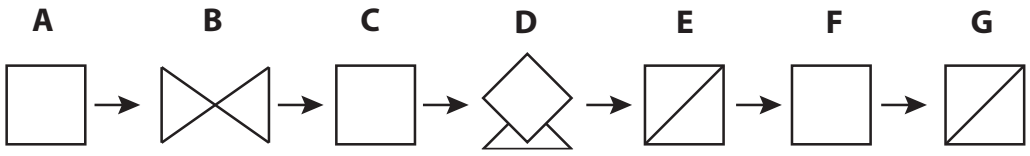

Figure 4.2 Summary of unit operations involved in the analysis of buffer-related variables in cell disruption studies. (A) Preparation of suspension at 4°C, relevant variables are the concentration of biomass in suspension ([X]), the concentration of salt ([NaCl)], the addition of DTT and the addition of glycerol. (B) High-pressure homogenisation at 10°C, relevant variables are the operating pressure and the number of passes. (C) Addition of non-ionic detergent following a one-hour incubation at 4°C, relevant operating variables are the type of detergent and the concentration of detergent. Crude lysate samples are taken after this incubation. (D) Centrifugation of crude lysate at 15000g, 4°C for 30 minutes. (E) 0.45µm & 0.22µm dead-end filtration of supernatant. (F) Incubation of supernatant at 4°C, investigated variable is the incubation time, samples are taken

beyond this point for particle size distribution analysis. (G) 0.22µm dead-end filtration of incubated sample. Samples were taken after incubations to study the effect of aggregation on product recovery.

4.3.2 Investigating the effect of detergent type and quantity

This set of experiments focussed on the effect on the solubility of Tandem Core protein in clarified lysate (step F in Figure 4.2) resulting from the addition of different types of detergents in lysate immediately after homogenisation (step C in Figure 4.2 at 1200 bar for 3 passes). Note that this addition step was after homogenisation, rather than before, as adding detergents during the homogenisation process could result in the formation of liquid-gas interfaces, typically detrimental to product quality. Detergent types used were Triton X100 (TX100), polysorbate 80 (PS80) and Tween 20 at final lysate concentrations of 1% and 0.1%. To capture the full effect of detergent-mediated solubilisation, an incubation period of 1 hour was performed prior to sample analysis. Total product and protein released in crude lysate was assessed through Western blotting with densitometry analysis (sections 2.5.9 and 2.5.11) and total protein analysis (section 2.5.5) on material after incubation in step C in Figure 4.2. Using similar methodology, any precipitation effects were monitored by passing incubated material through a second 0.22um (step G in Figure 4.2) filter and monitoring any decrease in protein concentration.

Looking at Figure 4.3A, after 72 hours of incubation, there is very little difference in core-positive material concentrations amongst different buffer-types. Note that Figure 4.3A depicts a compilation of densitometry data from two membranes, with

two separate two-point calibration curves. The first membrane contained buffer conditions at volumetric concentrations of 1%, whereas the second membrane depicts data from buffer concentrations of 0.1%. One must consider that this method of analysis is hard to reproduce quantitatively, and so it is not recommended that the two concentration-groups be compared directly. In addition, the difference in total soluble protein content after incubation was roughly the same in all buffer conditions, with marginally higher soluble protein content when using higher concentrations of TX100. Therefore, assuming that total soluble protein content would positively correlate to solubility of tHBc, and bearing in mind that total protein measurement data was more robust than membrane densitometry, TX100 was chosen as the detergent for future studies.

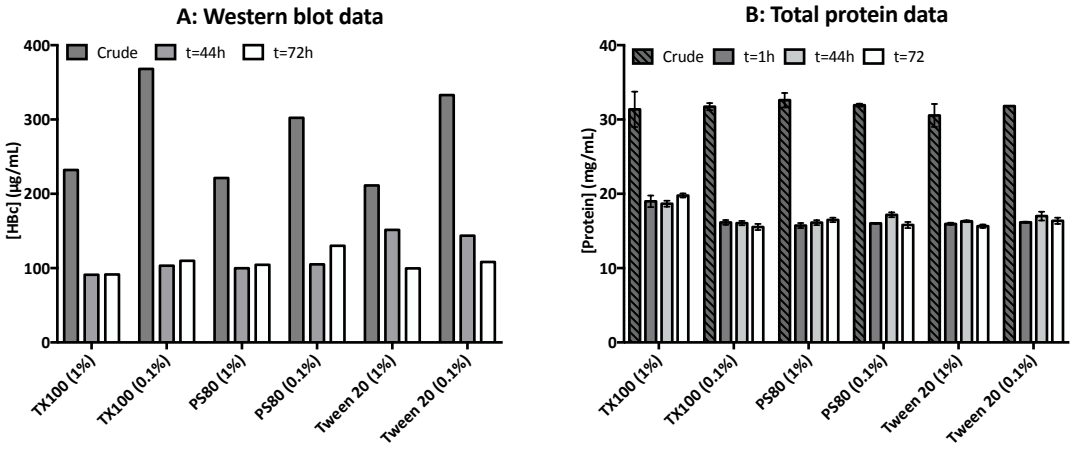

Figure 4.3 Effects of aggregation in clarified lysates after additions of various detergents at various concentrations. (A) Product recovery data in crude lysate and in clarified lysate after 44 hours and 72 hours of incubation at 4°C. Peak area was determined through densitometry analysis on Western blots. Peak area was converted to HBc concentration using a two-point calibration curve. No replicate data was generated. (B) Total protein content in crude lysate and in clarified lysate after 1 hour, 44 hours and 72 hours of incubation at 4°C. Error bars represent triplicate measurements taken using Nanodrop (section 2.5.5).

4.3.3 Buffer characterisation using a DoE methodology

The previous section described how several detergents contributed to buffer stability, product recovery and total cell disruption.

However, one unaddressed variable, contributing to the aforementioned key quality attributes, is the concentration of salt. Sources report that salt concentration can have a significant effect on the recovery of VLPs in storage (Shi *et al.*, 2005).

To study this effect in combination with the previous studies, an experiment was conducted as follows: after determining the most suitable detergent, a 4-factor, two-level, full factorial experimental screening design with 3 midpoints was executed. The factors chosen and their corresponding values are listed in Table 4-1.

Table 4-1 Experimental overview for a 4-factor, two-level, full factorial experimental screening design with 3 midpoints to characterise homogenisation conditions.

STD	[NaCl] (mM)	[TX100] (%vol.)	WCW (g/L)	t (h)
1	500	0.1	150	1
2	0	0.1	30	72
3	500	1.5	150	72
4	500	1.5	30	72
5	0	1.5	30	1
6	250	0.8	90	45
7	500	1.5	150	1
8	0	15	150	1
9	500	1.5	30	1
10	0	1.5	30	72
11	0	0.1	150	1
12	500	0.1	30	1
13	250	0.8	90	45
14	0	0.1	30	1
15	0	1.5	150	72
16	0	0.1	150	72
17	250	0.8	90	45
18	500	0.1	30	72
19	500	0.1	150	72

Several responses were chosen to generate screening results. Firstly, total soluble protein released (*R (mg/mL)*) was chosen to represent an overall level of cell disruption. As the solid loading percentage was varied, second variable was introduced to indicate a level of cell disruption per unit of biomass: specific protein release (R_s *(mg/gWCW*)). This measure is relevant when wanting to make efficient use of biomass input. Thirdly, overall product recovery was measured ([HBc] (µg/mL)). To generate a specific response for total protein release, a specific measure of product release was investigated*: HBc_s (µg/gWCW)).*

Screening models were generated with a minimum resolution of 6 to estimate all possible first-degree polynomial interactions.

The relative contribution of an individual factor was defined as the total of the sum of squares (Type III) of each factor-associated first-degree polynomial term as a percentage of the total of the sum of squares of all terms in the screening models. Note that quadratic terms were not included in the analysis of variance as these terms are aliased to all studied factors. Table 4-2 shows the level of overall response variance created by each factor in this screening experiment when all possible first-degree polynomial interactions are considered. The left side of Table 4-3 shows the goodness of fit of these models, using r^2. The robustness of these models is then assessed using r^2 *adjusted*. The same table subsequently shows formulae for each response model showing only significant model terms ($p<0.05$). Similarly, r^2 and r^2 *adjusted* of these simplified models are shown on the right-hand side of the table.

For total cell disruption/total protein extraction *(R)*, the amount of biomass in suspension *([X])* was found to be the most highly contributing and significant factor. This is unsurprising, as biomass is the sole source of protein in the experiment; no biomass results in no protein solubilising; and more biomass results in higher levels of released soluble protein, hence the positive gradient seem in the formula in Table 4-3. From Table II it can also be said that the contribution of the concentration of detergent is also relevant. This can be explained by the significant factor interaction existing between biomass concentration and detergent concentration seen in Table 4-3. The adjusted r^2 for both the complex model and simplified model are relatively high and close to their corresponding r^2 values. This shows that both models are relatively robust, not only having a high goodness of fit but also high predictive ability.

Table 4-2 shows that relative factor contributions for the response for specific cell disruption (R_s) differed significantly from the response *R*. Like with *R*, major response variance contributions are due to variation of *[X]* and *[TX]*. However, a combined 28.54% of response variance was caused by variance of *[NaCl]* and *t*. But because no significant model terms were associated with these factors, despite considerably contributing to model variance, a notable reduction in goodness of fit is seen in simplified models (Table 4-3). Moreover, the most significant factor in both responses for *R* and R_S, the concentration of biomass *([X]),* has the opposite effect. Maximising *R* requires maximising *[X]*, as indicated by the positive gradient in the simplified model. However, as indicated by the negative gradient in the simplified model term for R_s, maximising specific cell disruption requires minimising

[X]. The negative correlation between R_s and *[X]* can be explained by the fact that in the scenario with relatively less biomass, more kinetic energy directed to cell disruption is available per unit of biomass.

As seen in Table 4-2, the contribution of *[X]* is highest in the response model for total product release *([HBc]).* The concentration of salt *([NaCl])*, however, has much more of an effect on the release of product than the release of overall protein. The negative effect of salt concentration on product recovery can be explained by the effect of product "salting out" at higher concentrations of salt and relatively hydrophobic molecules.

Variance in HBc_s levels are mainly caused by variance in *[TX]*, followed by *[X]* and *[NaCl].* Table 4-3 shows that the both the complex and simplified models have a very high goodness of fit and are both very robust. As seen with *[HBc],* the negative model terms for *[TX]* and *[NaCl]* could be explained by salting out effects. Finally, like specific cell disruption, *[X]* has a negative effect on specific product release.

Table 4-2 Relative factor contributions in a four-factor, two-level, half-fractional factorial screening designs. Six different screenings were performed: Total cell disruption (*R*); specific cell disruption (R_S); total product released (*[HBc]*); specific product released (HBc_s); and recovered product purity relative to total protein concentration (*P*).

	Factor	**[X] (g/L)**	**[NaCl] (mM)**	**[TX] (%)**	**t (h)**
	Range	*(30,150)*	*(0.1,1.5)*	*(0.1,10)*	*(1,72)*
% contribution	***R (mg/mL)***	84.78%	2.98%	8.35%	3.89%
	R_s (µg/gWCW)	41.99%	11.97%	29.47%	16.57%
	[HBc] (µg/mL)	68.56%	13.93%	14.27%	3.24%
	HBc_s (µg/gWCW)	36.52%	19.82%	38.26%	5.39%
	P (‰)	27.20%	11.17%	55.00%	6.63%

Table 4-3 Models and goodness of fit. Screened responses for total cell disruption (*R*); specific cell disruption (R_S); total product released (*[HBc]*); specific product released (HBc_s); and recovered product purity relative to total protein concentration (*P*). Goodness of fit is given on the left-hand side of the table ('All terms') using r^2 and adjusted r^2 for models in which all possible first-degree polynomial terms have been considered. Following this, are the formulae describing first-degree polynomials containing only significant model terms ($p<0.05$). Highlighted in bold are internal functions. This method of displaying formulae allows for easy distinction between significance of model terms by ranking them in order of gradient. Integrating the absolute experimental domains of each corresponding factor would unnecessarily distort this ranking order.

	All terms (Complex model)			Significant terms (Simplified model)	
	R^2	R^2 adj.	Formulae	R^2	R^2 adj.
R (mg/mL)	0.992	0.951	$R = 17.9 + 8.20f([\boldsymbol{X}]) + 2.03f([\boldsymbol{TX}])$	0.933	0.925
Rs (µg/gWCW)	0.980	0.878	$R_s = 240 - 73.4f([\boldsymbol{X}]) + 48.7f([\boldsymbol{TX}]) - 39.3f([\boldsymbol{X}] \cdot [\boldsymbol{TX}])$	0.657	0.589
[HBc] (µg/mL)	0.966	0.799	$[HBc] = 37.5 + 17.0f([\boldsymbol{X}]) - 7.30f([\boldsymbol{NaCl}]) - 5.06f([\boldsymbol{TX}]) - 4.77f([\boldsymbol{X}] \cdot [\boldsymbol{TX}])$	0.905	0.878
HBc_s (ng/gWCW)	0.999	0.994	$[HBc]_s = 493 - 165f([\boldsymbol{TX}]) + 163([\boldsymbol{X}] \cdot [\boldsymbol{TX}]) - 146f([\boldsymbol{NaCl}]) - 137f([\boldsymbol{X}]) + 97.8f([\boldsymbol{X}] \cdot [\boldsymbol{NaCl}]) - 65.0f([\boldsymbol{TX}] \cdot \boldsymbol{t})$	0.951	0.926
P (‰)	0.985	0.909	$P = 2.28 - 0.888f([\boldsymbol{TX}]) - 0.754f([\boldsymbol{X}] \cdot [\boldsymbol{TX}]) - 0.478f([\boldsymbol{NaCl}])$	0.836	0.803

$f([\boldsymbol{X}]) = \frac{([X]-90)}{60}$ $f([\boldsymbol{TX}]) = \frac{([TX]-0.8)}{0.7}$ $f([\boldsymbol{NaCl}]) = \frac{([X]-250)}{250}$ $f(\boldsymbol{t}) = \frac{([X]-36.5)}{35.5}$

Detergent concentration was found to have the greatest effect on product purity, (Table 4-2). As seen with *[HBc]* and HBc_s in Table 4-3, salt and detergent concentration both have a negative effect on product purity, suggesting that certain host cell proteins have a significantly higher solubility than the expressed heterologous protein found in the soluble fraction.

This section has presented the effects of a variety of buffer components on a variety of performance criteria.

The performance of various detergents was investigated. It was found that polysorbate 80 and Triton-X100 performed similarly. For future work, the choice was made to continue working with Triton-X100 as this detergent had been used in most experiments thus far.

Finally, within the ranges investigated, salt concentration and detergent concentration had mostly negative effects on the five studied performance criteria. It should be noted that this is a two-level screening experiment and the existence of optima within the investigated domains is very likely. The section therefore recommends that future experiments re-evaluate the ranges chosen.

4.4 Homogenisation characterisation

After investigating the effects on various key quality attributes by a variety of lysis buffer properties, the next objective was addressed: characterising high-pressure homogenisation.

This series of experiments investigated the number of passes (*N*), the operating pressure (*p*) and the amount of biomass in suspension (*[X]*). Table 4-4 shows the experimental design space used to generate surface response models. Homogenisation operating procedures were as described in section 2.4.3.

Table 4-4 Design space for a 3-factor, two-level, Central Composite Design (CCD) with on-face axial points (α=1) and three centre points.

STD	[X] (g_{WCW}/L)	p (bar)	N
1	50	300	1
2	50	300	5
3	50	1200	1
4	50	1200	5
5	100	300	1
6	100	300	5
7	100	1200	1
8	100	1200	5
9	50	750	3
10	100	750	3
11	75	300	3
12	75	1200	3
13	75	750	1
14	75	750	5
15	75	750	3
16	75	750	3

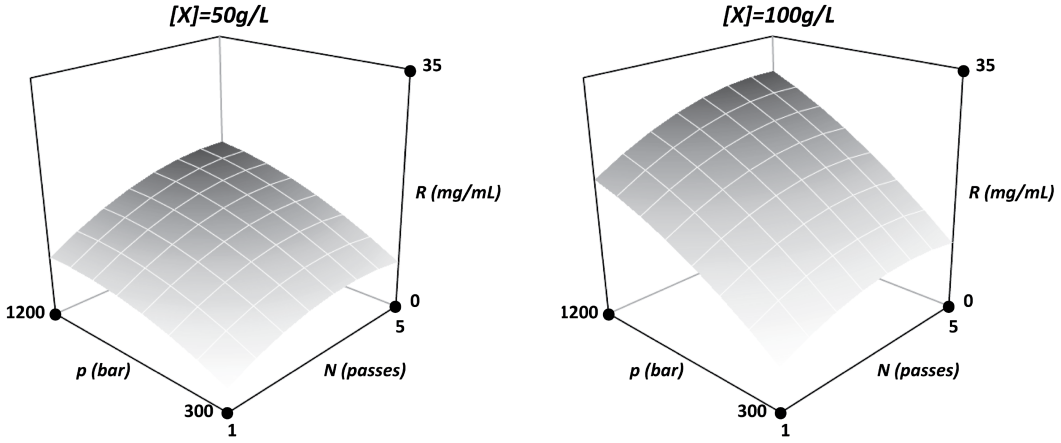

Figure 4.4 High Pressure Homogenisation Response Surface Models at different level of *[X]* for total cell disruption measured as total soluble protein, *R* (mg/mL*).*

Figure 4.4 shows the effect of biomass concentration *[X],* operating pressure *(p)* and the number of passes *(N)* on cell disruption *(R)* expressed as the total amount of soluble total protein (mg/mL). The response model can be described with the following formula: $R= -6.14\cdot10^{-1}N^2 -8.92\cdot10^{-6}\ p^2 + 1.38\cdot10^{-3}\ Np + 1.97\cdot10^{-4}\ p[X] + 4.20N + 8.4\cdot10^{-3}p -1.00\cdot10^{-3}[X] -5.15.$

As indicated by the term $1.97 \cdot 10^{-4}\ p[X]$ the factor interaction of biomass concentration and pressure has a positive effect on the level of cell disruption. The term is highly significant (p=0.00253), which is in contrast with the assumption made by others (Follows *et al.*, 1971), that the effect of biomass concentration on cell disruption is minor. This difference is most likely because Follows *et al.* investigated cell disruption at a significantly higher cell suspension concentration range (450-750g/L vs 50-100g/L) where the effect of biomass had diminished. Follows *et al.* used a different species of yeast, *Saccharomyces cerevisiae*, and the concentrations of biomass used are considerably higher than would be used for practical purposes since higher concentrations of cells lead to lower levels of specific cell disruption.

Unlike for pressure, the optimum number of passes seems to stay the same as the biomass concentration is changed. This is not surprising as no term for *[X]N* exists in the equation. However, the term $1.38 \cdot 10^{-3}\ Np$ indicates a positive factor-interaction between the number of passes and pressure, which is expected.

Finally, the two quadratic terms of N^2 and p^2 indicate optima for pressure and the number of passes, however it must be noted that these are quadratic estimates that are only valid within the investigated range. The response surface methodology in this series of experiments only considers linear or quadratic functions, not for instance, asymptotic functions which are very common in biological systems and bioprocesses. Therefore, supposed quadratic optima found at the end of the

experimental window, could also indicate asymptotic limits. However, establishing this would require further work and does not change the main output of this study.

It should also be noted that cell disruption (*R*) is just one measure of performance in the context of cell disruption. As mentioned previously, this paper aims to investigate five different performance criteria as response surfaces. Instead of displaying four additional response models, Figure 4.5 summarises the factor settings for various performance optima derived from corresponding surface models.

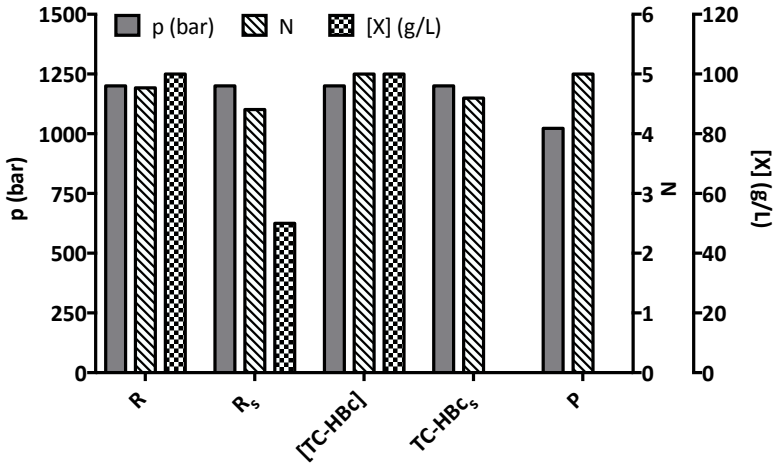

Figure 4.5 High Pressure Homogenisation performance optima summary. Factor values and optima for various responses: *R* (r^2=0.98), R_S (r^2=0.97), *[TC-HBc]* (r^2=0.92), $TC\text{-}HBc_S$ (r^2=0.90) and *P* (r^2=0.77).

Looking at Figure 4.5, one can conclude that different operating conditions are required to achieve different maximum performance levels. For instance, maximum pressure (1200 bar) is required to achieve maximum cell disruption (R). However sub-maximal pressure (1000 bar) is preferable for maximum purity (P). Note that in Figure 4.5 biomass is excluded from the performance maxima of the response

models for specific product recovery (TC-HBC_S) and purity (P) as this was found to have insignificant contributions to variance for these specific models ($p>0.05$).

In summary, this section has described the characterisation and optimisation of high-pressure Homogenisation on the basis of five different performance criteria. Different optimum operating conditions were established, depending on the investigated performance criterion.

4.5 Development of small-scale cell disruption methods

This section will explore a variety of cell disruption platforms suitable for small-scale ($V \leq 1mL$) use. However, this section will focus mainly on enzymatic cell disruption with lyticase and cell disruption using Adaptive Focussed Acoustics (AFA). Additional small-scale cell disruption methods are discussed in the appendix section.

4.5.1 Enzymatic lysis with lyticase

Various sources have reported that enzymes can be used to disrupt yeast cells by digesting cell wall components. Several sources report that the enzyme, lyticase, hydrolyses beta-glucans in fungal cell walls, resulting in the transformation of cells into so-called protoplasts (Burden, no date; Salazar and Asenjo, 2007; Miyajima *et al.*, 2009). Lacking cell walls, these protoplasts are much more susceptible to mechanical stress.

It was therefore hypothesised that an enzymatic pre-treatment step using lyticase could be effective in performing cell disruption for *P. pastoris*.

To assess the effect of incubation time with lyticase on harvested material, 30 g_{WCW}/L cells (from 3.2.1) was incubated in 20mM Tris-based lyticase buffer (2.4.6). Protein concentrations in crude lysate were measured every ten minutes using the Nanodrop method described in section 2.5.5. After 90 minutes, an addition of Triton-X100 was added to achieve a concentration of 0.1% (vol.), in order to maximise disruption and assess this augmentation effect of detergent on cell disruption.

It was estimated that maximum disruption through enzymatic lysis had occurred after around 20 minutes but with a higher degree of certainty after 60 minutes (see Figure 4.6). After 90 minutes of incubation, subsequent addition of TX100 resulted in further disruption of cells, reflected by an average 31.9% increase in measured protein concentration.

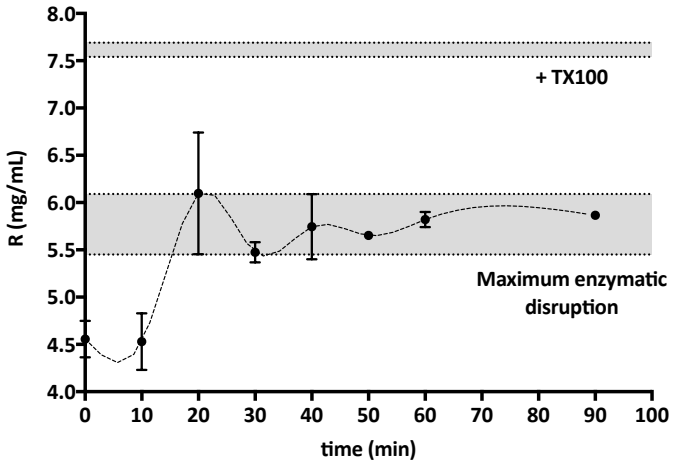

Figure 4.6 Enzymatic cell disruption over time of cells incubated with lyticase at a biomass concentration of 30g_{WCW}/L. Standard deviations per data point represent triplicate total protein measurements of crude lysate. The lower highlighted grey area represents the fluctuation of maximum enzymatic cell disruption and was defined as the overall standard deviation of all measurements taken between 20-90 minutes. The upper grey area represents the level of cell disruption after 90 minutes of incubation and an addition of Triton-X100 to achieve a concentration of 0.1% (vol.).

The efficacy of lyticase as a disruption agent was further assessed by incubating harvested material (from 3.3.1) at various cell densities in 20mM Tris-based lyticase buffer (2.4.6). Samples were stored overnight at 4°C prior to clarification as described in section 2.4.10. Total cell disruption, *R*, was quantified as total soluble protein release through using Nanodrop analysis, as described in section 2.5.5. The results of this experiment are shown in Figure 4.7.

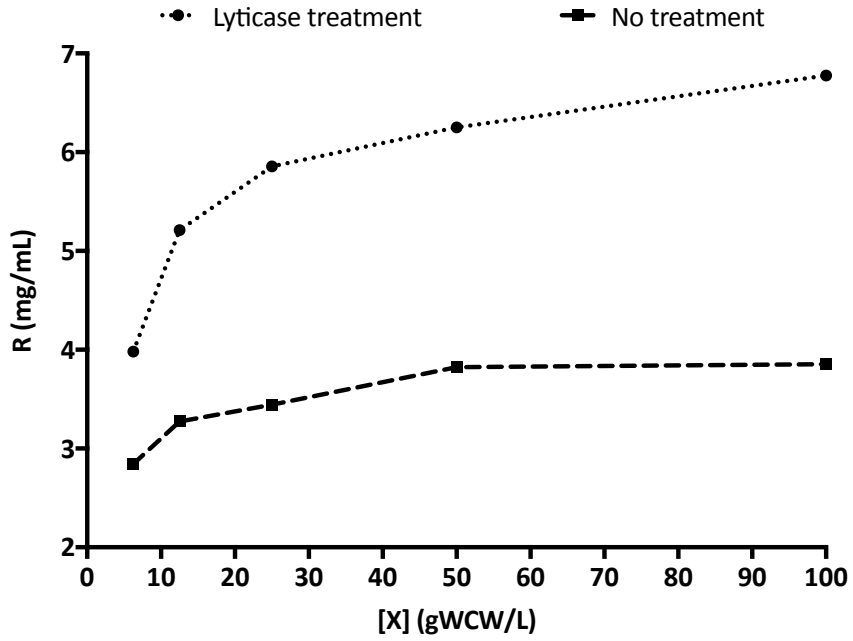

Figure 4.7 Total cell disruption of cells incubated with and without lyticase at various biomass concentrations. Cell disruption averages and standard deviations (not visible) are derived from triplicate measurements of total soluble protein in clarified lysate

Figure 4.7 shows significant increased levels of total of cell disruption, at a variety of biomass concentrations as a result of incubating with lyticase. It also seems that, as the concentration of biomass increases, the specific level of cell disruption, indicated by gradient, decreases. This indicates a substrate dependant process, typical of enzymatic reactions. Note that a similar trend is seen when no lyticase is applied. The absolute increase of observed cell disruption resulting from an

increase in biomass concentration can be attributed to various mechanisms including freeze-thaw-mediated lysis, autolysis and detergent-mediated lysis.

Furthermore, the use of lyticase also allows for the recovery of soluble core-positive material as shown in section 10.3.

4.5.2 Adaptive Focused Acoustics

Adaptive Focused Acoustics (AFA) has been shown to be a suitable miniaturised platform for the disruption of the yeast, *Saccharomyces cerevisiae* (Wenger *et al.*, 2008). An AFA device generates acoustic shock waves in the kilohertz (KHz) region. These sonic waves cause controlled cavitation at a focal point within the sample vessel (Covaris, 2011). The device can process up to 12-96 samples in a batch, enabling rapid, high-throughput, non-contact cell disruption. Figure 4.8 explains the operating parameters associated with the Covaris E210.

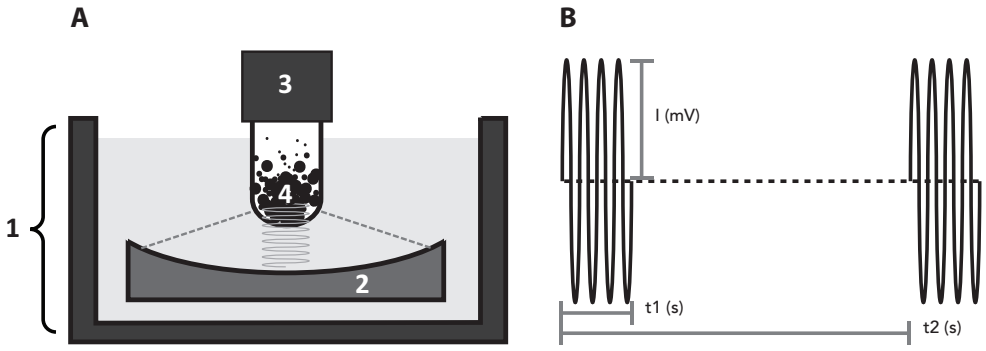

Figure 4.8 Covaris components and parameters. (A1) Tank containing cooled degassed deionised water, (A2) acoustic transducer, (A3) vial containing suspended cells and (A4) acoustic focal zone. (B) Covaris parameters: Duty Factor, *DF* (%), refers to the relative time between sonic bursts, hence $DF = \left({}^{t1}/_{t2}\right) \cdot 100\%$. Intensity, *I* (mV), denotes the amplitude of the sonic wave. Cycles per burst, *cpb*, refers to the amount of cycles per acoustic wave. In this case, *cpb* is 4. Time, *t* (s), simply refers to the total time a sample is exposed to sonication treatment. *[X]* (g/L) refers to the wet cell biomass concentration in suspension.

Process variables contributing to cell disruption can be described as follows:

$$\boldsymbol{R = f(DF, I, cpb, t, [X])} \qquad \textbf{Equation 4-3}$$

Duty Factor, *DF (%),* refers to the relative time between sonic bursts, hence

$$\boldsymbol{DF = \left({}^{t1}/_{t2}\right) \cdot 100} \qquad \textbf{Equation 4-4}$$

Intensity*, I (mV),* denotes the amplitude of the sonic wave. Cycles per burst, *cpb*, refers to the amount of cycles per acoustic wave. Time, *t (s),* refers to the total acoustic exposure time. *[X] (g/L)* refers to the wet cell biomass concentration in suspension.

Although there are studies showing promising results for extracting intracellular product from *S. cerevisiae*, no such extensive studies exist for AFA-mediated cell disruption of *P. pastoris*.

AFA-mediated cell disruption of *P. pastoris* was investigated by incubating harvested material (from 3.3.1) at various cell densities in 20mM Tris-based lyticase buffer (2.4.6). These samples were subsequently transferred to 6mL Chromacol tubes and exposed to AFA treatment for 0 or 120 seconds using maximum settings for Duty Factor, Cycles Per Burst and Intensity on Power Tracking mode. Samples were stored overnight storage at 4°C prior to clarification. Total cell disruption, *R*, was quantified as total soluble protein release through using Nanodrop analysis (section 2.5.5). The results of this experiment are shown in Figure 4.9.

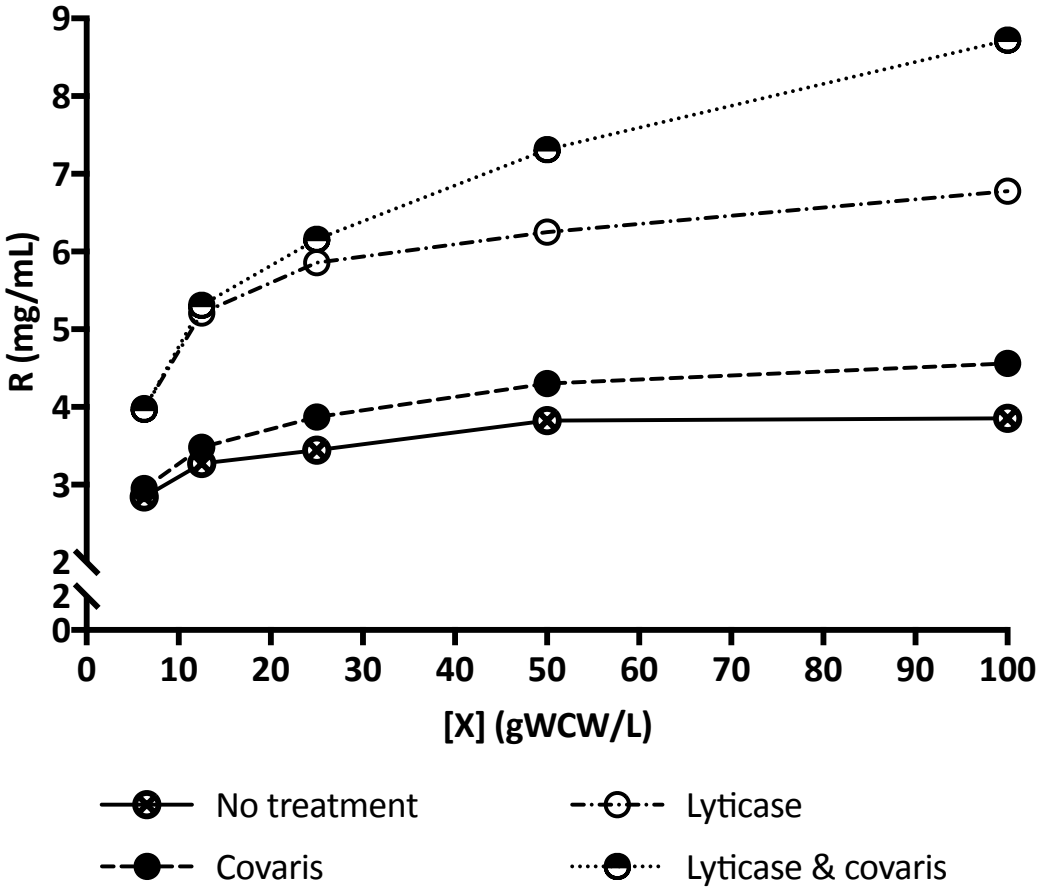

Figure 4.9 Total cell disruption through various methods. Cell disruption averages and standard deviations (not visible) are derived from triplicate measurements of total soluble protein in clarified lysate.

Figure 4.9 shows that AFA treatment resulted in higher levels of cell disruption than detergent treatment alone. In addition, it shows that lyticase pre-treatment augmented AFA-mediated cell disruption.

In addition to cell disruption, Figure 4.10 shows that both modes of AFA, with and without lyticase pre-treatment, result in recovery of soluble product.

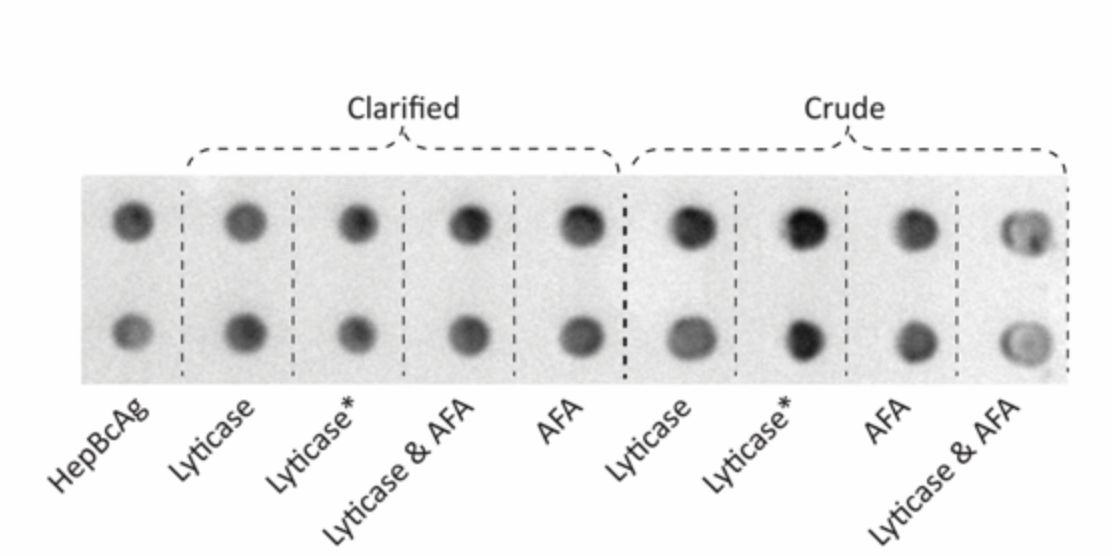

Figure 4.10 Qualitative dot blot analysis of crude and clarified lysates following lyticase treatment, AFA treatment and lyticase treatment followed by AFA treatment at a biomass concentration of 100gWCW/L and (*) 50gWCW/L. Measurements were performed in duplicate. Reference material (hepatitis B core antigen) was used as a positive control.

Due to these results and the automated, non-contact character of the platform, it was found that this platform was the best candidate for follow-up research out of all the small-scale cell disruption platforms investigated.

4.6 Adaptive Focused Acoustics characterisation

The previous sections described how Adaptive Focused Acoustics was selected as the best candidate for a small-scale cell disruption platform.

This section introduces a formalised approach on how this platform was further optimised through the following steps:

- Screening operating settings
- Comparing vessel designs
- Investigating the use implementation of an enzymatic pre-treatment step

In addition to this, it should be noted that unless stated otherwise, milliTUBEs were used for all experiments, 20mM TRIS buffer was substituted for 50mM MOPS buffer and all harvested material was produced by the fermentation described in section 7.4.1.

4.6.1 Adaptive Focused Acoustics Screening

In order to eliminate insignificant variables, an AFA screening was performed. The Covaris was operated as described in section 2.4.5, unless stated otherwise.

Table 4-5 shows the results of the first screening experiment which was a five-factor, half fractorial design was used to generate two factor screenings.

Table 4-5 Relative factor contributions in a five factor, two-level, half fractional factorial screening designs. Two different screenings were performed: One for *R* and one for R_s.

	Factor	DF (%)	cpb	I (mv)	t (s)	[X] (g/L)
	Range	(0.1,20)	(50,1000)	(0.1,10)	(60,300)	(10,100)
% contribution	Total disruption (R)	23.5%	8.6%	9.9%	8.7%	49.1%
	Specific disruption (R_S)	1.6%	0.1%	0.3%	0.3%	97.7%

Biomass *([X])* was found to be the most significant factor for both specific (R_s) and total cell disruption (*R*). However, the magnitude of the range of *[X]* may have diminished the significance of other factors and may therefore not be representative.

Time, on the other hand, initially seems to be insignificant relative to other factors which was in contradiction to other sources (Wenger *et al.*, 2008). This could be because the effect of time had already approached a limit and is not significant within the studied range.

To account for these two points of discussion, another screening was performed. In this four-factor full fractional screening, biomass was fixed at *[X]*=55g/L so responses R and R_S were equal. Additionally, the cycles per burst were found to be insignificant. Therefore, in further experiments this was fixed at 1000 cpb. The exposure times were lowered, as it was hypothesised that longer exposure times were excessive to achieving disruption and would therefore not show significant contribution to response variance. The results of this screening are shown in Table 4-6. Following two screening experiments it was found that biomass, duty factor, intensity and acoustic exposure time were significant factors. Increasing acoustic exposure time did not have a significant effect beyond 60s so this factor was fixed at this value, while other significant factors were varied to generate response surface models.

Table 4-6 Relative factor contributions in a four-factor, two-level, full-fractional factorial screening design

	Factor	*DF (%)*	*cpb*	*I (mv)*	*t (s)*
	Range	(0.1,20)	(50,1000)	(0.1,10)	(10,60)
% contribution	***Total cell disruption (R=R_S)***	40.1%	0.4%	38.4%	21.1%

4.6.2 Adaptive Focused Acoustics Optimisation

Figure 4.12 shows the effects of different types of vials, milliTUBE and Chromacol vials (shown in Figure 4.11), on cell disruption performance. The milliTUBE vials consistently outperformed Chromacol vials for every response. Like homogenisation, different settings are required depending on the chosen response maxima.

For total cell disruption and specific cell disruption (see Figure 4.12A) maximum performance values of the Chromacol configuration were found to be slightly lower than that for milliTUBE vials. However, the values for maximum product recovery and specific recovery (Figure 4.12B) are significantly lower in the Chromacol tubes when compared to milliTUBE vials. This in turn, results in minimal product purity relative to the milliTUBE configuration since P=[TC-HBc]/R.

Figure 4.11: Different tube types used in adaptive focused acoustics optimisation: 6mL Chromacol tube (left) and 1mL MilliTube (right).

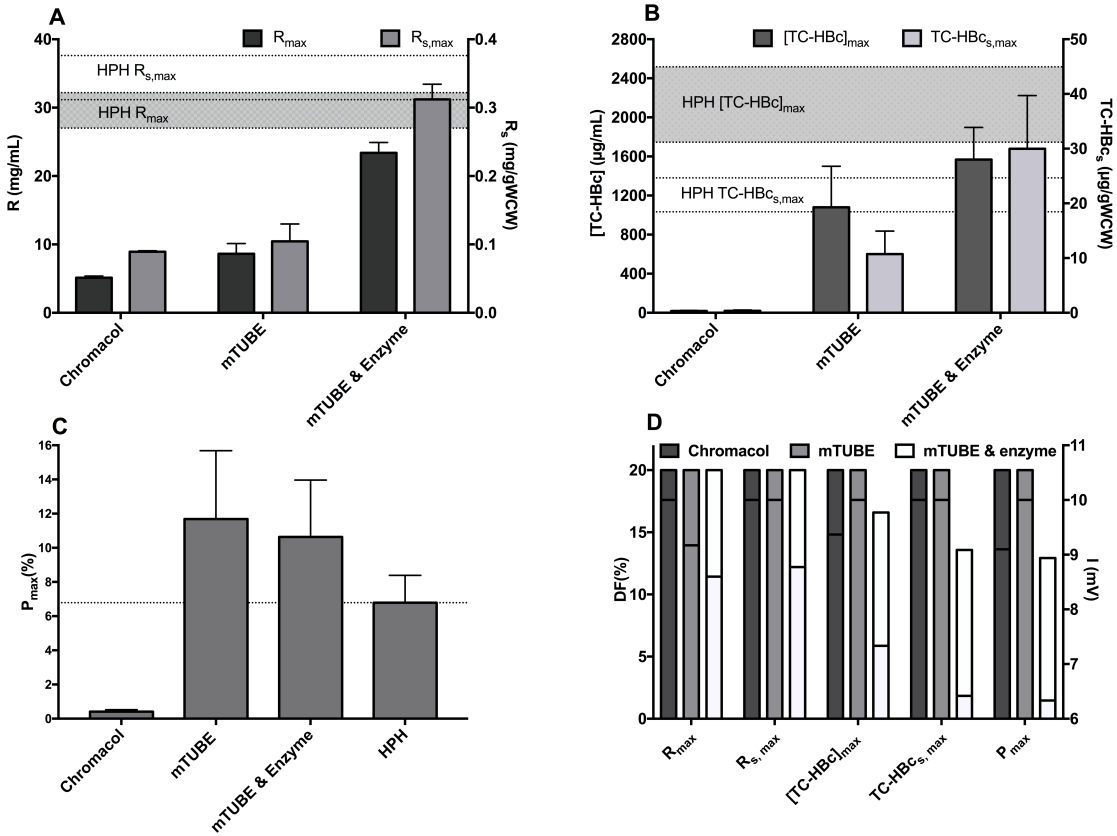

Figure 4.12 Summary of Response Surface Models for Adaptive Focused Acoustic methods. (A-C) Maximum response ranges for various performance criteria. (A) On the left y-axis, the level of total cell disruption (*R*); on the right y-axis, the specific level of cell disruption (R_s). (B) On the left y-axis the level of total product recovery *([TC-HBc]*); on the right y-axis, the specific level of product recovery (*TC-HBc*$_s$). (A, B) The horizontal grid lines display the maximum model ranges of the studied responses of High Pressure Homogenisation (HPH). (C) Purity (*P*) performance ranges. (D) Different factor values to achieve optima. Intensity values are superimposed over duty factor values.

Figure 4.13, showing Response Surface Models of total cell disruption (R) using milliTUBE and Chromacol vials, explains these differences in overall performance. Although AFA-mediated disruption using the Chromacol vials does influence disruption within a certain range, as proven by significant contribution to variance by the studies factors, its effect on overall disruption is most likely diminished by other mechanisms, such as cell autolysis during fermentation, osmotic shock, mechanical stress or heat lysis during buffer resuspension. This is demonstrated in Figure 4.13 by a relatively flat response model compared to that of the milliTUBE configuration.

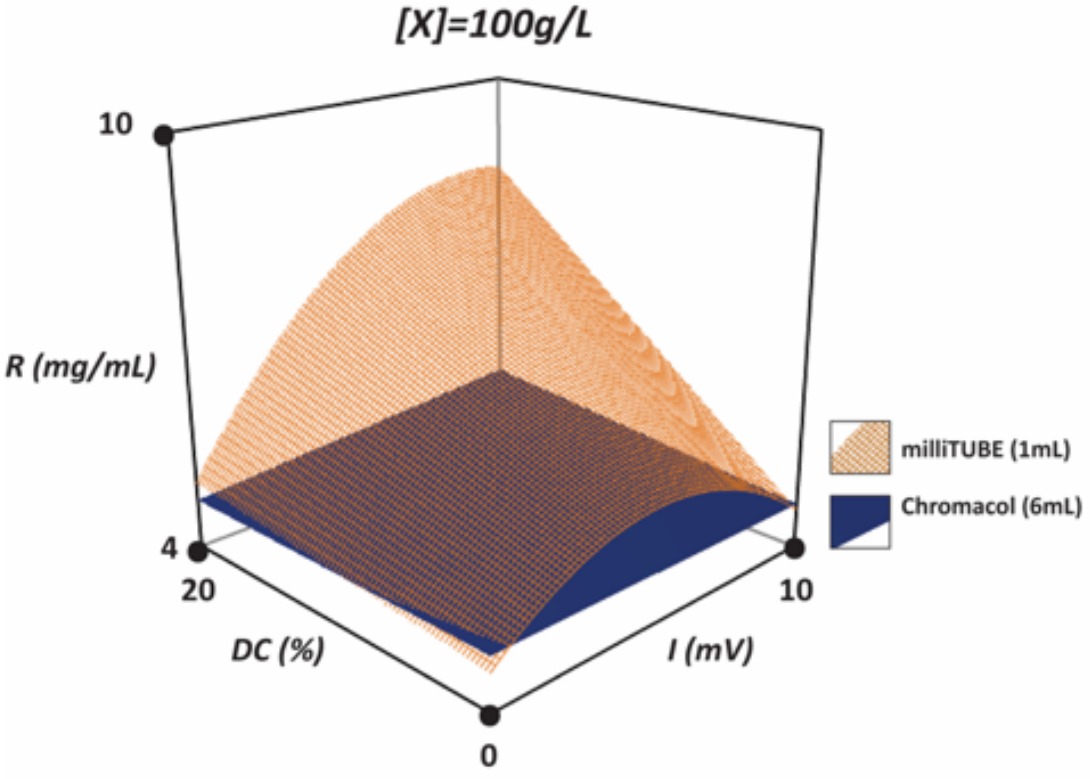

Figure 4.13: Effect of vial design. Response Surface Models at *[X]*=100g/L WCW and *t*=60s for two configurations of AFA-mediated cell disruption. (1) Using 6mL Chromacol vials: (2) 1mL milliTUBE vials with integrated acoustic fibres.

The underlying mechanism of these differences relates to total vessel volume, total working volume and additional mechanical stress through the addition of an AFA fiber in the milliTUBE vials. Because the total vessel volume of Chromacol vial is significantly larger than that of a milliTUBE vial, acoustic energy is dispersed over a larger volume resulting in lower levels of disruption. Additionally, the Chromacol vials used a 17% working volume, whereas milliTUBE vials were filled to almost 100%. This difference in relative working volumes was a result of standardising the experiments to absolute working volumes of 1mL. Larger relative working volumes lead to more efficient cell disruption processes as acoustic energy is directed less to mixing and forming liquid-air interfaces and more towards cavitation nucleation in the lysis medium. As higher working volumes reduce mixing in the sample, the added integrated AFA fiber facilitates the generation of high numbers of uniformly distributed cavitation bubbles. The simultaneous collapse of these high-energy liquid-gas interfaces results in higher levels of mechanical energy and thus higher levels of cell disruption (Covaris Inc., 2017).

Several sources report that the enzyme, lyticase, hydrolyses beta-glucans found in yeast and fungal cell walls form so-called protoplasts (Burden; Miyajima et al. 2009; Salazar & Asenjo 2007). Lacking cell walls, these protoplasts are much more susceptible to mechanical stress. Therefore, cell suspension samples were pre-treated with lyticase to augment subsequent AFA-mediated cell disruption.

Results in Figure 4.11 show that this enzymatic pre-treatment step resulted in significantly higher overall performance in AFA-mediated cell disruption compared to the other small-scale methods lacking this enzymatic treatment step.

Anand et al. (2007) described how pre-treatment of *Escherichia coli* cells with guanidium hydrochloride, affected the first order disruption rate constant, k, for homogenisation. Li et al. (2012) subsequently described this rate constant in the context of adaptive focused acoustic in the following equation at fixed biomass concentration:

$$\boldsymbol{ln\left(\frac{R_m - R_0}{R_m - R}\right) = kt} \quad \textbf{Equation 4-5}$$

Where R_m is the total maximum available amount of protein available for release. R_0 is the level of protein release prior AFA-mediated cell disruption and was found to be 4.17mg/mL for untreated cells and 16.90mg/mL for enzymatically pre-treated cells ([X]=100g/L). R is the observed level of disruption and t is the acoustic exposure time fixed at 60s. No absolute value for R_m was obtained experimentally but as our purposes are comparative, R_m was chosen to be the maximum observed level of protein release from homogenisation at *[X]=100g/L* (R_m*=29.6mg/mL*).

It was expected that the lyticase pre-treatment step would increase the disruption rate constant due to the weakening of cells. However, because of this it was difficult to determine how much of the observed cell disruption was due to non-mechanical lysis, such as osmotic shock, or due to AFA-mediated cell disruption. Using the AFA milliTUBE configuration, we found that the disruption rate constant for AFA

mediated cell disruption preceded by lyticase treatment ($k_{LY,AFA}$= $6.82 \cdot 10^{-3}$ s^{-1}) was 2.65 times higher than the disruption rate constant of AFA-mediated cell disruption (k_{AFA}=$2.57 \cdot 10^{-3}$ s^{-1}). This strongly suggests that enzymatic pre-treatment augments performance of AFA-mediated cell disruption by weakening cells.

Because of this pre-treatment, maximum levels of AFA-mediated cell disruption almost matched maximum levels achievable with homogenisation. Moreover, the performance of the augmented method for maximum specific product recovery was superior to that of homogenisation. Likewise, both platforms using milliTUBE achieved higher product purity levels than High Pressure Homogenisation.

4.7 Causes and effects of micronised debris

The previous section demonstrated the differences in various response between High Pressure Homogenisation and various modes of adaptive focused acoustics.

This section aims to explain the mechanisms involved behind the performance of the discussed cell disruption methods.

Figure 4.14A shows Particle Size Distribution (PSD) curves of various unclarified lysates from disrupted samples. The amount of micro debris (less than 1µm) generated through High Pressure Homogenisation is much higher than that generated through AFA-mediated disruption. These high levels of micro debris are associated with high levels of cell disruption, however also impede recovery of

product. This is demonstrated in Figure 4.14B where homogenisation is shown to lead to lower product purities than AFA despite higher levels of cell disruption.

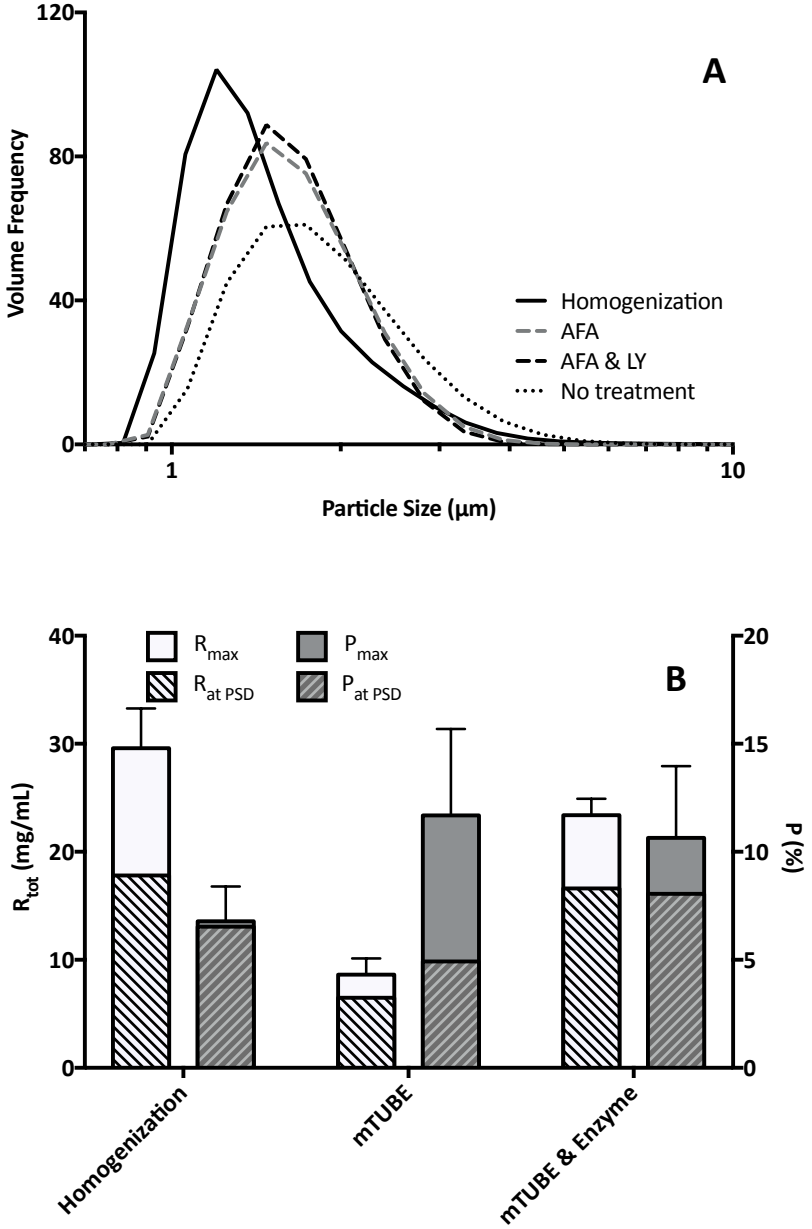

Figure 4.14: Effect of micronised debris on purity. (A) PSD analysis derived from static light scattering analysis (2.5.12), of untreated cells in suspension *([X]*=50g/L) and respective crude lysates disrupted by homogenisation (*p*=1200 bar, *N*=5 passes, *T*=8±2°C), AFA (*I*=10mV, *DF*=20%, *cpb*=1000, *t*=60s, *T*=9±1°C) and AFA preceded by lyticase incubation (*t*=1hr, *T*=25°C). (B) Modelled quantitative analysis of clarified lysates disrupted by AFA, AFA preceded by lyticase incubation and homogenisation. Two levels of cell disruption (*R*) and product purity (*P*) are given: the maximum values of the respective Response Surface Models (R_{max}, P_{max}) and, superimposed, the levels of disruption and purity ($R_{at\ PSD}$, $P_{at\ PSD}$) achieved at the same disruption conditions represented by the PSD curves. Error bars correspond to performance maxima only.

Interestingly, the enzymatic pre-treatment step does not significantly lead to a change in PSD even though it significantly increases levels of cell disruption. This is probably why the maximum purity level of milliTUBE AFA-mediated cell disruption is comparable to that of milliTUBE AFA-mediated cell disruption preceded by lyticase treatment.

The generation of micro debris during homogenisation can lead to subsequent formation of aggregates as seen Figure 4.15A and Figure 4.15B. The fluorescent BODIPY stain shows neutral lipids in intact cells that are adhered with cell debris. Such aggregation can severely limit the release of the hydrophilic product into the supernatant upon subsequent clarification.

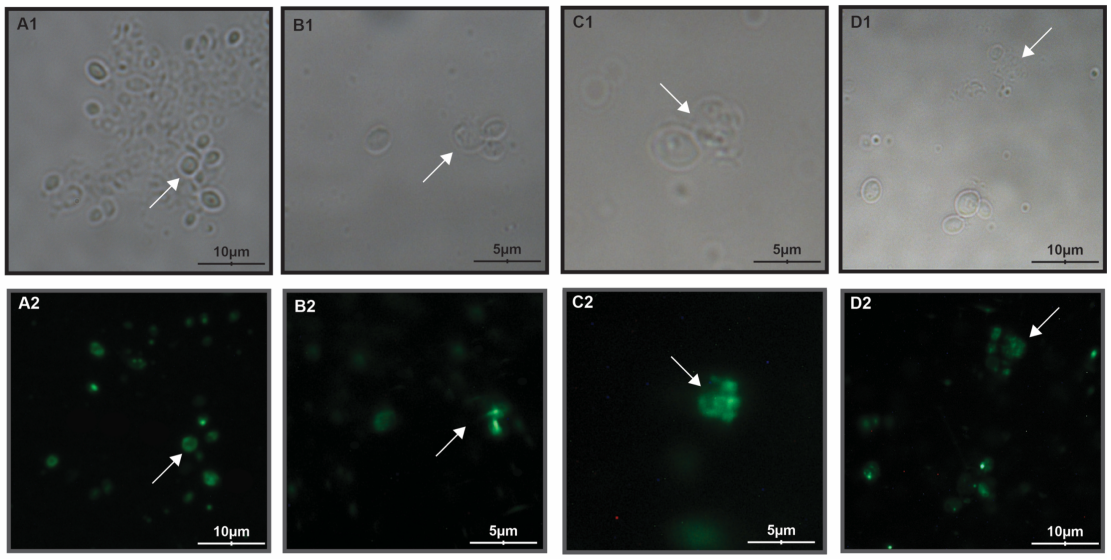

Figure 4.15: Microscopy images of lysates (*[X]*=50g/L) at 101.5x magnification. (A1) Bright light microscopy image of homogenised sample (*p*=1200 bar, *N*=5 passes, *T*=8±2°C) and (A2) corresponding fluorescent image; arrows indicate intact cell. (B1, B2) AFA (*I*=10mV, *DF*=20%, *cpb*=1000, *t*=60s, *T*=9±1°C); arrows indicate disrupted cell. (C1, C2) Lyticase pre-treated sample; arrows show protoplasted cell. (D1, D2) Lyticase pre-treated sample followed by AFA mediated disruption (*I*=10mV, *DF*=20%, *cpb*=1000, *t*=60s, *T*=9±1°C).

In contrast, AFA mediated cell disruption does not seem to generate micronised debris. As shown in Figure 4.15B, a cell disrupted by AFA, visible with bright light microscopy yet undetectable using BODIPY staining, has released its lipid-positive content out of the cell wall through a single cavity, without generating large amounts of micronised debris. Figure 4.15C shows what appears to be an enzymatically treated cell that has lost its characteristic shape, most likely due to cell wall disintegration by lyticase. Figure 4.15D shows a similar effect on enzymatically treated cell subsequently disrupted using AFA.

4.8 Process scaling and cell disruption mimicry

This chapter has made many comparisons between the relatively large-scale industry standard of cell disruption, High Pressure Homogenisation, and various modes of small-scale, non-contact Adaptive Focused Acoustics methods.

For some performance criteria, it was found that it is possible for AFA to outperform High Pressure Homogenisation.

However, when scaling down a process, the goal is often not to outperform the large-scale surrogate, but to mimic its performance as closely as possible. To achieve this, one must define these performance criteria.

This section considered a scenario where the objective was to maximise product purity from a homogenisation process and, using enzyme-augmented AFA, mimic

both the level of purity *and* the corresponding level of product recovery at the same homogenisation parameter settings.

First, operating ranges were defined by determining standard deviation windows of (1) maximum product purity, P, and (2) the corresponding level of product recovery, *[TC-HBc,]* at maximum purity of High Pressure Homogenisation.

Overlaying the purity and product recovery response functions for enzymatically augmented AFA ('AUG. AFA)', at the specified ranges, yielded the operating domains for duty factor and intensity settings. The corresponding operating window is displayed in white in Figure 4.16.

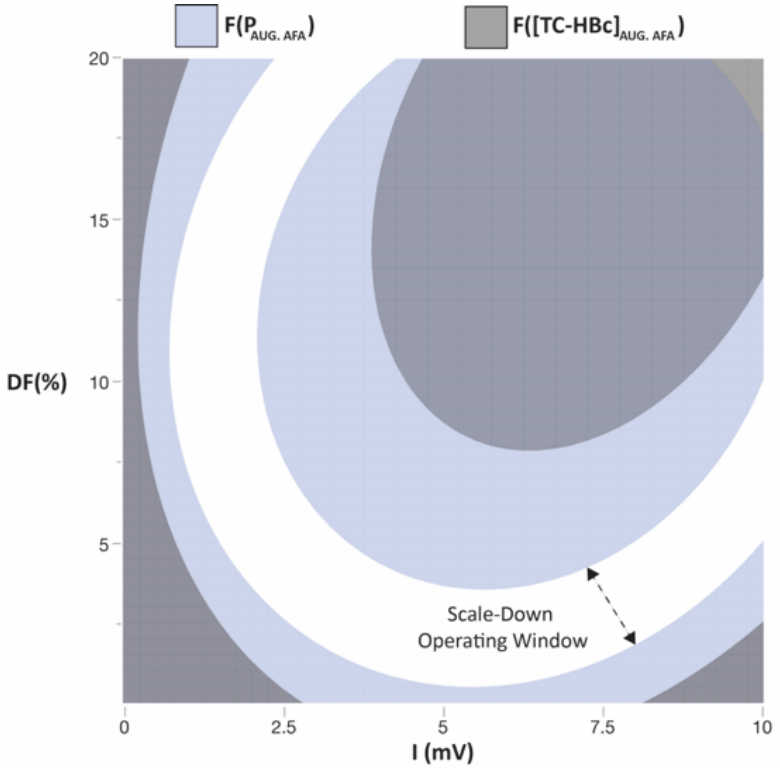

Figure 4.16: Microscale process mimicry. Overlay plot of enzymatically augmented AFA-mediated cell disruption funcions of purity, $F(P_{AUG.\ AFA})$, and product recovery, $F([TC\text{-}HBc]_{AUG.AFA})$ at $[X]$=50 g/L and t=60s.

It should be noted that the above scenario serves as an example of how the proposed scale-down methods can be implemented to mimic any large-scale cell disruption process. By assigning multiple performance criteria, more defined conditions can be achieved to mimic performance at small-scale. Likewise, the same methodology can be applied to scale-up a micro-scale cell disruption process. This suggests that the scale-down AFA-mediated cell disruption can be used in a wide array of applications, such as high-throughput buffer development and micro-scale fermentation sampling whilst retaining scalable significance.

4.9 Summary

This chapter presents the development of a high-performance, high-throughput, non-contact, automated, small-scale, scalable disruption tool for microbial bioprocess development. This was achieved through several steps.

Firstly, the effects of lysis buffer components were investigated to establish standardised buffer conditions amongst different cell disruption platforms. This was followed by a full characterisation of high-pressure homogenisation to establish scale-down performance targets. After this, various small-scale disruption cell disruption techniques were compared. It was found that enzymatic lysis with the enzyme lyticase and Adaptive Focused Acoustics (AFA) provided the most suitable characteristics for further investigation. Following this, screening and subsequent optimisation experiments were conducted on AFA. These investigations were carried out through statistical screening methods, the development of twenty

Response Surface Models, various modes of microscopy and particle size distribution analysis.

Significant process parameters for AFA were found to be acoustic exposure time, biomass suspension concentration, duty factor and intensity. After performing optimisation experiments for these significant factors, it was found that the design of the vessel in which samples were sonicated had a great impact on all cell disruption performance criteria. Specifically, it was found that the use of milliTUBE vials greatly enhanced performance compared to using Chromacol vials. Performance was subsequently augmented using a lyticase pre-treatment step. This augmentation led to matching and even outperforming homogenisation performance, depending on the performance criterion investigated. The resulting overlap of performance ranges between HPH and AFA-mediated cell disruption allowed for small-scale AFA performance mimicry of High Pressure Homogenisation. Performance mimicry was enhanced by matching multiple performance criteria, as opposed to a single criterion, using overlay plot analysis.

Chapter 5 Microscale upstream process platform development

5.1 Introduction

Chapter 3 addressed the need for miniaturised fermentation platforms to study upstream process-associated factors. The same chapter also identified a challenge in the fact that existing small-scale extraction methods were either poorly scalable, non-automatable, unsuitable for high-throughput sample processing or ineffective in recovering significant levels of soluble heterologous product.

Chapter 4 addressed this challenge through the development of high-performance, high-throughput, non-contact, automated, small-scale, scalable disruption tools for microbial bioprocess development.

Having developed the necessary tools to implement in miniaturised upstream process development, the use of miniaturised fermentation platforms can be investigated to address several key observations made in chapter 3.

5.2 Aim and objectives

The aim of this chapter is to develop a miniaturised fermentation platform intended for rapid upstream process characterisation and to use this platform to determine critical process parameters for rapid upstream process characterisation. Much work has already been conducted for *E. coli* growth in microtitre plates, but literature is

limited for *P. pastoris* cultivation and protein expression (Islam *et al.*, 2008; Kensy, Engelbrecht and Büchs, 2009a). The specific objectives are therefore to:

- Assess growth kinetics in microtitre plates to determine their capacity of sustaining biomass for *P. pastoris* cultures.
- Investigate bolus feeding induction strategies for high density *P. pastoris* cultures
- Investigate alternative non-methanol feeding strategies

5.2.1 Batch-mode growth kinetics screening in microtitre plates

The most basic form of a miniaturised fermentation characterisation platform is typically a 24 deep-well plate. To maximise gas transfer, square-shaped wells are frequently used. This is because the corners of the square wells disturb the flow in a way similar to that of baffles in shake flasks or fermenters (Duetz *et al.*, 2000).

The most critical stage of upstream processing of Tandem Core VLPs expressed by *P. pastoris* is the induction phase. However, before this platform could be used to investigate induction phase-associated factors, the capacity of these deep well plates to sustain biomass, and mimic batch-mode fermentation, had to be investigated.

This section will hence aim to describe the capacity of microtitre plates to sustain growth according to the specific growth rate, μ (h^{-1}). This growth rate can be described by the equation:

$$\frac{d[X]}{dt} = \mu[X] \qquad \textbf{Equation 5-1}$$

Where:

- *[X]* is the concentration of biomass
- *t* is the time (hours)
- μ is the specific growth rate (per hour)

The amount of biomass at any given time, *t*, can be described as:

$$[X]_t = [X]_0 e^{\mu t} \qquad \textbf{Equation 5-2}$$

Hence:

$$\ln[X]_t = \ln[X]_0 + \mu t$$

$$\therefore \mu = \frac{\ln[X]_t - \ln[X]_0}{t}$$

Therefore, to study this response, $[X]_0$ and *t*, needed to be varied.

In addition, the working volume, V_w, was also varied. The working volume directly affects the Oxygen Transfer Rate. This in turn has a significant effect on cellular metabolism. This effect is explained in more detail in 10.4.

In this way, the working volume, V_W, could affect (1) the consumption of substrate, (2) the expression of product and (3) the production of biomass. It was therefore hypothesised that an increase in working volume would result in a decrease in culture growth.

This was investigated through a three-factor, two-level, full-fractional factorial screening designs. Microtitre plate batch-mode fermentations were carried out as described in section 2.3.1.

A screening model, consisting of 8 experimental points and 3 centre points, was generated with a minimum resolution of 5 to estimate all possible first-degree polynomial interactions. The design space and corresponding response values are shown in Table 10-2.

The relative contribution of an individual factor was defined as the total of the sum of squares (Type III) of each factor-associated first-degree polynomial term as a percentage of the total of the sum of squares of all terms in the screening models. Note that quadratic terms were not included in the analysis of variance as these terms are aliased to all studied factors.

Table 5-1 shows the level of overall response variance created by each factor in this screening experiment when all possible first-degree polynomial interactions are considered. The left side of Table5-2 shows the goodness of fit of these models, using r^2. The robustness of these models is then assessed using r^2 *adjusted*. The same table subsequently shows formulae for each response model showing only significant model terms ($p<0.05$). Similarly, r^2 and r^2 *adjusted* of these simplified models are shown on the right-hand side of this table.

As seen in Table 5-1 each studied factor has different levels of relative variance contribution depending on the response investigated. Table 5-2 shows that r^2

consistently decreases as the models are simplified through elimination of insignificant factors; and shows significant model terms.

Table 5-1: Relative factor contributions in a three-factor, two-level, full-fractional factorial screening designs

Factor		$[X]_0$ (OD_{600})	V_w (mL)	t (h)
Range		(20,100)	(0.5, 2.5)	(2,30)
% contribution	μ (h^{-1})	3.51%	44.4%	52.1%

Table 5-2: Model and goodness of fit for specific growth rate (μ). Goodness of fit is given on the left-hand side of the table ('All terms') using r^2 and adjusted r^2 for the model in which all possible first-degree polynomial terms have been considered. Following this, is the formula describing first-degree polynomials containing only significant model terms ($p<0.05$). Highlighted in bold are internal functions[11]

	All terms (Complex model)			Significant terms (Simplified model)	
	R^2	R^2 adj.	Formula	R^2	R^2 adj.
μ (h^{-1})	0.948	0.826	$\mu = 4.91 \cdot 10^{-2} - 3.78 \cdot 10^{-2} f(\boldsymbol{t}) - 3.37 \cdot 10^{-2} f(\boldsymbol{V_W}) + 3.35 \cdot 10^{-2} f(\boldsymbol{V_W} \cdot \boldsymbol{t})$	0.904	0.863

$$f([\boldsymbol{X_0}]) = \frac{[X_0] - 60}{40} \qquad f(\boldsymbol{t}) = \frac{t - 16}{14} \qquad f(\boldsymbol{V_W}) = \boldsymbol{V_W} - 1.5$$

Table 5-2 shows that the specific growth rate, μ, decreases as the incubation time increases. This is most likely due to nutrient depletion over time. The same response also has a negative correlation with the working volume. This proves the hypothesis that an increase in working volume would result in a decrease in culture growth.

Overall, model simplification for this response resulted in an increase in model robustness, represented by an increase in *adjusted* r^2. Interestingly, the initial

[11] This approach to displaying formulae allows for easy distinction between significance of model terms by ranking them in order of gradient. Integrating the absolute experimental domains of each corresponding factor will unnecessarily distort this ranking order.

concentration of biomass was not found to have a significant effect on specific growth rate within the chosen range. However, this could be due to oxygen-limiting and substrate-limiting effects.

The specific growth rate was of interest as one of the aims of this section was to determine the capacity of this fermentation platform to sustain biomass. For a fermentation platform to be deemed suitable for maintaining or growing biomass at any particular time, one could define that the specific growth rate needs to be equal to or higher than 0 (μ≥0). Therefore, the factor exclusion of $[X]_0$ in the response for specific growth rate presented an opportunity to establish a generalised correlation between V_W and t when μ=0. Such a correlation would allow for the determination of harvest times at various working volume, and vice-versa, in which no overall culture death had occurred.

To investigate for how long at what working volume the wells were able to sustain biomass (μ=0), a correlation was established between V_W and t:

$$\therefore \boldsymbol{t}_{(\mu=0)} = \frac{-4.47}{\boldsymbol{V_W} - 2.63} + 30.14$$

Equation 5-3

The calculation of this correlation is shown in section 10.4.3 the appendix. A graphic depiction of this correlation is shown in Figure 5.1.

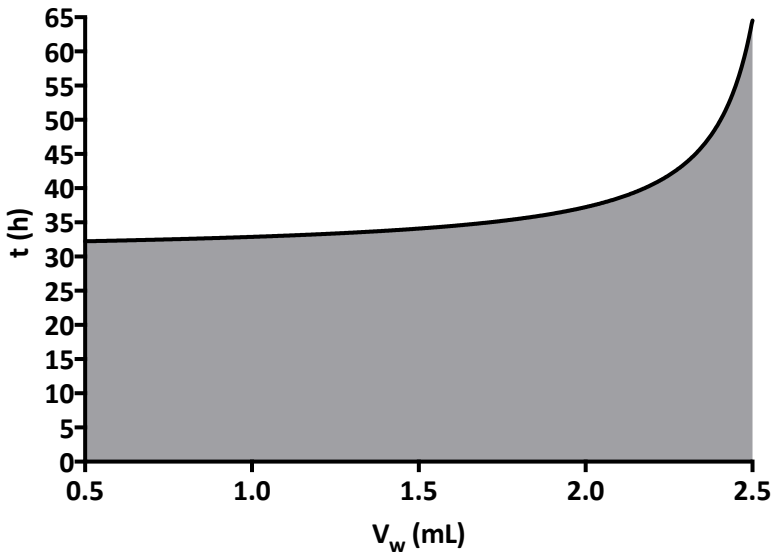

Figure 5.1 Correlation between V_W and t when μ=0 in BMGY media within the range of working volume investigated.

This asymptotic correlation, within the ranges investigated, shows the maximum incubation time at various working volumes. For coordinates outside the grey area in Figure 5.1, the specific growth rate is negative, indicating cell death due to substrate depletion. As postulated earlier, increasing the working volume reduces the concentration of dissolved oxygen in the media. When oxygen becomes limited, substrate consumption rates decrease. This results in extended batch-mode fermentation times at decreased average growth rates. Ideally, this correlation should be verified outside the investigated ranges in future work.

Nonetheless, microtitre plates can be deemed a suitable platform for cultivating *P. pastoris* and, as shown in table 10-2, can sustain cell densities of up to 140 OD_{600}.

5.3 Assessing methanol toxicity

With the aim of creating a microtitre platform, characterising VLP assembly in an upstream bioprocess environment, the next objective was to assess whether this platform could be used as a screening tool to study the induction phase, in which methanol is introduced.

The introduction of methanol into culture media presents a two-fold challenge: too low concentrations of methanol would result in the starvation of the cultures due to a lack of carbon source, whereas too high concentrations of methanol would result in culture death due to toxicity. As mentioned in chapter 1, in typical bioreactor-based fermentations of *P. pastoris*, this issue is usually mitigated through the use of continuous feeding of methanol, often combined with sophisticated feedback control systems. However, in microtitre based platforms, no such control mechanisms are available. Therefore, a major conceptual challenge in the implementation of microtitre plate-based fermentations in upstream process development for *P. pastoris*, is translating continuous feeding into bolus additions of methanol.

To choose appropriate ranges in subsequent screening experiments, the effect of methanol concentration was studies on both Mut^S and Mut^+ cultures. The protocol for this is described in section 2.3.1.

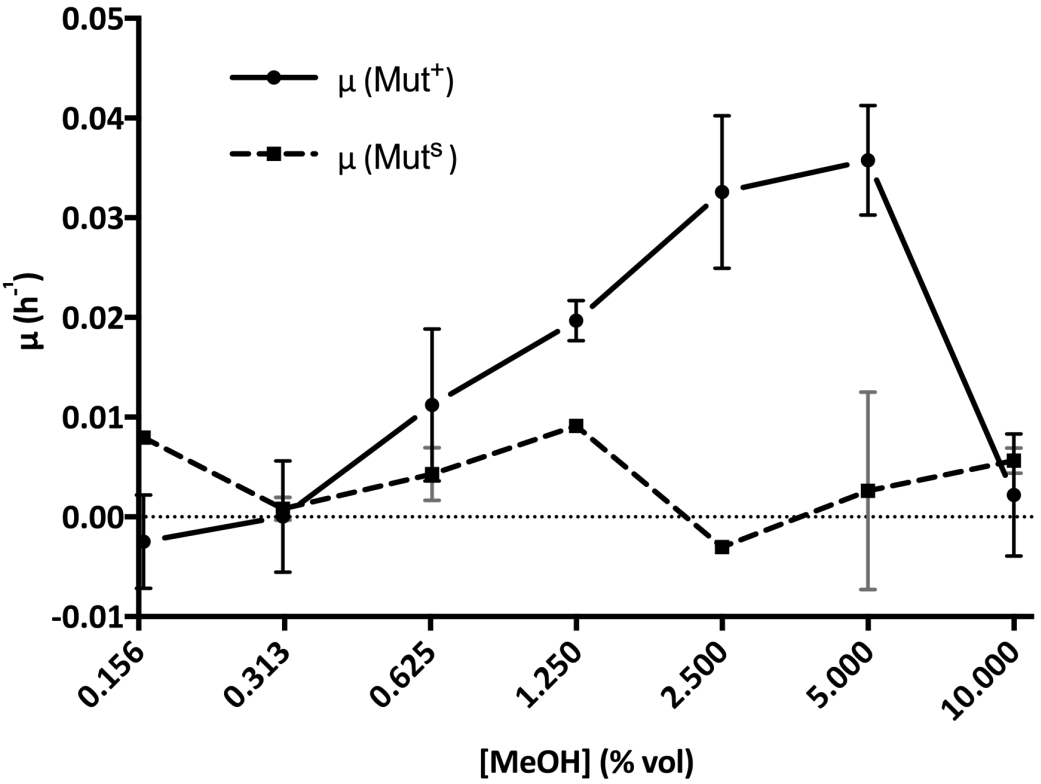

Figure 5.2 Effect of methanol concentration on growth rates after 24 hours of induction of *P. pastoris* Mut^S and Mut$^+$ cultures. Starting cultures were standardised to OD_{600}=8; and standard deviations and averages were derived from duplicate measurements.

As seen in Figure 5.2, for the Mut$^+$ phenotype, it seems that a minimum methanol concentration of 0.313 % vol. is required to avoid culture starvation. The growth rate decreases between 5.00 and 10.0% vol. methanol. Negative specific growth rates were observed at the latter concentration, indicating toxicity.

Note that the error in some measurements its quite high, obscuring trends. This error could have possibly been introduced by differences in media evaporation between wells. The effect of methanol on the Mut^S phenotype is less conclusive, but seems to not introduce the same level of variance as with the Mut$^+$ phenotype.

In addition, methanol toxicity seems to occur at lower ranges than the Mut^+ phenotype, between 1.25 and 2.50% vol. methanol.

However, both profiles, assuming a One-Factor-at-a-Time (OFAT) approach, do not take variance of time and initial biomass into account. The additions of these variables will be investigated in the next section using a DoE-based screening approach.

5.4 Induction screening in microtitre plates

Methanol toxicity profiles were assessed for both *P. pastoris* phenotypes in section 5.3. Understanding the differences in the manifestation of methanol toxicity between the two studied phenotypes allowed for suitable range determination in subsequent microtitre plate-based screening experiments for induction phase-associated factors.

5.4.1 Initial induction screening results

The previous sections have investigated (1) the capacity of microtitre plates to sustain biomass and mimic bioreactor batch-mode performance and (2) the effects of methanol concentrations on the growth of both phenotypes of *P. pastoris*.

This section will perform screening experiments to determine whether microtitre plates are a suitable platform to study induction processes for *P. pastoris* cultures.

In addition to the specific growth rate, studied in section 5.2.1, this section will therefore also investigate (1) the total level of expression of tHBc protein, *[tHBc]* (ng/mL) and (2) the specific expression of tHBc protein, $tHBc_s$ (ng/mL/AU).

These aims were examined through a three-factor, two-level, full-fractional factorial screening designs. Microtitre plate fermentations were carried out as described in section 2.3.1. Primary recovery was performed as described in sections 2.4.1, 2.4.6 and 2.4.10. Product analysis was performed with dot blots (2.5.10 and 2.5.11).

The design space and corresponding response values are shown in Table 10-3 for the Mut^+ phenotype and in Table 10-4 for the Mut^S phenotype. Based on the methanol toxicity profiles shown in Figure 5.2, the methanol concentrations chosen for the Mut^+ phenotype were higher than for the Mut^S phenotype.

The relative contribution of an individual factor was determined as described in 5.2.1. Table 5-3 and Table 5-4 show the level of overall response variance created by each factor in screening experiments for Mut^+ and Mut^S phenotypes respectively. The left side of Table 5-5 and Table 5-6- Mut^+ and Mut^S phenotypes respectively- show the goodness of fit of these models, using r^2. The robustness of these models was subsequently assessed using r^2 *adjusted*. The same tables subsequently show formulae for each response model showing only significant model terms ($p<0.05$). Similarly, r^2 and r^2 *adjusted* of these simplified models are shown on the right-hand side of these tables.

Table 5-3: Relative factor contributions in a three-factor, two-level, full-fractional factorial screening designs investigating induction of *P. pastoris* Mut$^+$ with pure methanol in microtitre plates: Specific growth rate (μ) total heterologous protein expression ([tHBc]), specific heterologous protein expression ($tHBc_S$), initial biomass concentration ($[X]_0$), initial methanol concentration ([MeOH]) and induction time (t). *: Curvature detected in screening response model.

Factor		$[X]_0$ (OD_{600})	[MeOH] (% vol.))	t (h)
Range		(18, 90)	(0.5, 5.0)	(2,30)
% contribution	μ (h^{-1})*	32.3%	0.748%	67.0%
	[tHBc]* (μg/mL	21.9%	5.30%	72.8%
	$tHBc_s$ (μg/mL/OD_{600})	58.6%	5.91%	35.5%

Table 5-4: Similar to Table 5-3, this table shows the relative factor contributions of a screening design investigating induction of *P. pastoris* Mut^S with pure methanol in microtitre plates.

Factor		$[X]_0$ (OD_{600})	[MeOH] (% vol.)	t (h)
Range		(18, 90)	(0.3, 1.0)	(2,30)
% contribution	μ (h^{-1})*	47.4%	0.282%	52.3%
	[tHBc]* (μg/mL)	48.0%	3.91%	48.0%
	$tHBc_s$ (μg/mL/OD_{600})	46.3%	7.38%	46.3%

Table 5-5: Models and goodness of fit for screening responses investigating induction of *P. pastoris* Mut$^+$ with pure methanol in microtitre plates: Specific growth rate (μ) total heterologous protein expression ([tHBc]), specific heterologous protein expression ($tHBc_S$), initial biomass concentration ($[X]_0$), initial methanol concentration ([MeOH]) and induction time (t). Goodness of fit is given on the left-hand side of the table ('All terms') using r^2 and adjusted r^2 for models in which all possible first-degree polynomial terms have been considered. Following this, are the formulae describing first-degree polynomials containing only significant model terms (p<0.05).

	All terms (Complex model)			Significant terms (Simplified model)	
	R^2	R^2 adj.	Formulae	R^2	R^2 adj.
μ (h^{-1})	0.885	0.618	$\mu = -0.147 + 0.170\, f(\boldsymbol{t}) - 0.0970 f([\boldsymbol{X_0}] \cdot \boldsymbol{t}) + 0.0939([\boldsymbol{X_0}])$	0.877	0.825
[tHBc] (μg/mL)	0.734	0.113	$[\boldsymbol{tHBc}] = 53.9 + 23.7 f(\boldsymbol{t})$	0.514	0.460
$tHBc_S$ (μg/mL/OD_{600})	0.995	0.984	$\boldsymbol{tHBc_S} = 2.89 - 2.25 f([\boldsymbol{X_0}]) + 1.45 f(\boldsymbol{t}) - 1.14 f([\boldsymbol{X_0}] \cdot \boldsymbol{t})$	0.921	0.888

$$f([\boldsymbol{X_0}]) = \frac{[X_0] - 54}{36} \qquad f(\boldsymbol{t}) = \frac{t - 16}{14}$$

Table 5-6: Similar to Table 5-5, this table shows the models and goodness of fit for screening responses investigating induction of *P. pastoris* Mut^S.

	All terms (Complex model)			Significant terms (Simplified model)	
	R^2	R^2 adj.	Formulae	R^2	R^2 adj.
μ (h^{-1})	0.935	0.782	$\mu = -0.119 + 0.131f(\boldsymbol{t}) - 0.121f([\boldsymbol{X_0}] \cdot \boldsymbol{t}) + 0.119f([\boldsymbol{X_0}])$	0.931	0.902
[tHBc] (μg/mL)	0.921	0.736	$[\boldsymbol{tHBc}] = 2.07 + 2.84f(\boldsymbol{t}) + 2.84f([\boldsymbol{X_0}] \cdot \boldsymbol{t}) + 2.84f([\boldsymbol{X_0}])$	0.871	0.816
[tHBc] (μg/mL/OD_{600})	0.924	0.749	$\boldsymbol{HBc_S} = 0.0415 + 0.0570f(\boldsymbol{t}) + 0.0570f([\boldsymbol{X_0}] \cdot \boldsymbol{t}) + 0.0570f([\boldsymbol{X_0}])$	0.829	0.756

$$f([\boldsymbol{X_0}]) = \frac{[X_0] - 54}{36} \qquad f(\boldsymbol{t}) = \frac{t - 16}{14}$$

In the case of both phenotypes, the concentration of methanol seems to have little effect on any of the responses. However, this seems to be conflicting the data shown in Figure 5.2, where the concentration of methanol shows to have a clear effect on the growth rate. There are several possible explanations for this. By looking at the measured concentrations of hepatitis B core-positive material Table 10-4 one could conclude that the lack of observed expression suggests that the design space was not appropriately chosen for Mut^S phenotype. For both phenotypes, many responses models indicated curvature. Also noteworthy is the fact that, contrary to the Mut^S cultures, expression of tHBc was consistently observed in each run for the Mut$^+$ phenotype (Table 10-3) However, the screening methodology described here excludes, by definition, the implementation of quadratic terms in response models since it is not possible to attribute this curvature to a single factor. The omission of the hypothetically significant quadratic model terms could explain the observed poor goodness of fit for μ and *[tHBc].* It is highly likely that the effect of methanol concentration can be described as solely quadratic since, as discussed earlier, relatively low concentrations result in culture

starvation, whereas too high concentrations result in toxic conditions. However, the same could be hypothesised for time, *t*, and initial biomass concentration, *($[X_0]$).* Regarding the hypothesis of the existence of an optimum for initial biomass concentration, one could imagine that the quadratic relationship of the ratio of biomass concentration to methanol concentration is highly significant: relatively high ratios would result in relatively rapid substrate depletion followed by culture stagnation, while relatively low ratios could result in the culture to be overwhelmed by methanol resulting in cell death. An optimum for harvesting time could be explained by the fact that *P. pastoris,* like many microorganisms, is known to require a minimum period of adaptation to metabolically adjust to the introduction of a different carbon source (Invitrogen Corporation, 2002), this affecting all of the studied responses, whilst too long incubation times would, again, result in culture stagnation due to depletion of carbon source.

5.4.2 Second induction screening

Results from the previous section (5.4.1) suggested to investigate different ranges for the design space for the Mut^S phenotype induction screening. Therefore, a second screening was conducted in which the following hypotheses were addressed:

The only observed instances of expression were found in runs in which both initial biomass concentrations and incubation time were relatively high (Table 10-4). This effect could be a result of a relatively long adaptation period to methanol in this fermentation platform. Therefore, an increase in both concentration of biomass and

incubation time could yield higher expression for the Mut^S phenotype induced in microtitre plates.

- If the above strategy indeed would improve the levels of heterologous protein expression, then it should also be ensured that no carbon source depletion occurs after the adaptation period to methanol. Therefore, a slightly higher and a wider range of methanol concentration – accommodating for higher levels of biomass and incubation times- could yield better expression of tHBc for the Mut^S phenotype induced in microtitre plates.

To further investigate the limitations for the design space for the Mut^+ phenotype the following hypotheses were formulated:

- The previous section (5.4.1) showed good expression of heterologous protein for the Mut^+ phenotype. In addition, all the observed growth rates were negative. This, just like with the Mut^S phenotype this could be due to the requirement of relatively long adaptation times to methanol. Also, it was shown that time had a positive effect on specific growth rate, volumetric expression and specific expression. Hence, an increase in incubation times could yield higher expression for the Mut^+ phenotype induced in microtitre plates.

- As harvest time increases, lower cell density cultures were shown to express volumetrically more product than higher cell density cultures. This is probably because higher cell density cultures require higher amounts of substrate (MeOH) for cell maintenance. Hence, lowering the initial biomass concentration at induction could yield higher expression for the Mut^+ phenotype induced in microtitre plates.
- The existence of toxic methanol concentrations for Mut^+ phenotypes has previously been established (Figure 5.2) at fixed initial biomass concentrations and incubation times. Section 5.4.1 partially aimed to study how this toxicity was affected through variation of initial biomass concentrations and incubation times. However, no obvious toxic effects were observed. To establish a screening range limitation for methanol, the range of methanol concentration needs to be widened until toxic effects are observed.

Table 10-5 and Table 10-6 show the second set of screenings' design spaces and corresponding response values for the Mut^+ and Mut^S phenotype respectively and Table 5-8 show the corresponding analysis of variance and Table 5-9 and Table 5-10 show the corresponding screening models.

Table 5-7: Relative factor contributions in second three-factor, two-level, full-fractional factorial screening designs investigating induction of *P. pastoris* Mut$^+$ with pure methanol in microtitre plates: Specific growth rate (μ) total heterologous protein expression ([tHBc]), specific heterologous protein expression ($tHBc_S$), initial biomass concentration ($[X]_0$), initial methanol concentration ([MeOH]) and induction time (t).

Factor		$[X]_0$ (OD_{600})	[MeOH] (% vol.))	t (h)
Range		(9, 45)	(2.5, 7.5)	(16,46)
% contribution	μ (h^{-1})	24.74%	48.78%	26.48%
	[tHBc] (μg/mL	41.23%	26.61%	32.16%
	$tHBc_s$ (μg/mL/OD_{600})	35.25%	31.89%	32.86%

Table 5-8: Similar to Table 5-7, this table shows the relative factor contributions of a screening design investigating induction of *P. pastoris* Mut^S with pure methanol in microtitre plates. *: Curvature detected in screening response model.

Factor		$[X]_0$ (OD_{600})	[MeOH] (% vol.)	t (h)
Range		(90, 135)	(0.5, 2.5)	(16,46)
% contribution	μ (h^{-1})	51.87%	22.54%	25.60%
	[tHBc] (μg/mL)*	63.41%	30.94%	5.65%
	$tHBc_s$ (μg/mL/OD_{600})	71.75%	4.84%	23.41%

Table 5-9: Models and goodness of fit for second screening designs investigating induction of *P. pastoris* Mut$^+$ in microtitre plates: Specific growth rate (μ) total heterologous protein expression ([tHBc]), specific heterologous protein expression ($tHBc_S$), initial biomass concentration ($[X]_0$), initial methanol concentration ([MeOH]) and induction time (t). Goodness of fit is given on the left-hand side of the table ('All terms') using r^2 and adjusted r^2 for models in which all possible first-degree polynomial terms have been considered.

	All terms (Complex model)			Significant terms (Simplified model)	
	R^2	R^2 adj.	Formulae	R^2	R^2 adj.
μ (h^{-1})	0.993	0.978	$\mu = 0.0185 - 0.0134f([\boldsymbol{MeOH}]) + 0.0126f([\boldsymbol{MeOH}] \cdot \boldsymbol{t}) + 0.0118f([\boldsymbol{MeOH}] \cdot [\boldsymbol{X_0}]) - 0.0107f([\boldsymbol{MeOH}] \cdot [\boldsymbol{X_0}] \cdot \boldsymbol{t}) - 0.00684f([\boldsymbol{X_0}] \cdot \boldsymbol{t})$	0.993	0.986
[tHBc] (μg/mL)	0.956	0.852	$[\boldsymbol{tHBc}] = 0.430 + 0.306f([\boldsymbol{X_0}] \cdot \boldsymbol{t})$	0.251	0.168
[tHBc] (μg/mL/OD_{600})	0.961	0.869	$\boldsymbol{HBc_S} = 0.0431 - 0.0541f([\boldsymbol{X_0}])$	0.163	0.067

$$f([\boldsymbol{X_0}]) = \frac{[X_0] - 27}{18} \qquad f([\boldsymbol{MeOH}]) = \frac{[MeOH] - 5}{2.5} \qquad f(\boldsymbol{t}) = \frac{t - 31}{15}$$

Table 5-10: Similar to Table 5-9, this table shows the models and goodness of fit for screening responses investigating induction of *P. pastoris* Mut^S.

	All terms (Complex model)			Significant terms (Simplified model)	
	R^2	R^2 adj.	Formulae	R^2	R^2 adj.
μ (h^{-1})	0.989	0.963	$\mu = 5.92 \cdot 10^{-4} - 3.76 \cdot 10^{-3} f([\boldsymbol{X_0}]) + 2.65 \cdot 10^{-3} f([\boldsymbol{X_0}] \cdot \boldsymbol{t}) + 2.53 \cdot 10^{-3} f([\boldsymbol{MeOH}]) - 1.63 \cdot 10^{-3} f([\boldsymbol{MeOH}] \cdot \boldsymbol{t}) - 8.81 \cdot 10^{-4} f(\boldsymbol{t})$	0.982	0.963
[tHBc] (μg/mL)	0.833	0.442	$[\boldsymbol{tHBc}] = 8.30 - 1.52 f([\boldsymbol{X_0}]) + 1.40 f([\boldsymbol{MeOH}] \cdot [\boldsymbol{X_0}]) -$	0.754	0.692
[tHBc] (μg/mL/OD_{600})	0.939	0.797	$\boldsymbol{HBc_S} = 0.0761 - 0.0235 f([\boldsymbol{X_0}]) + 0.0128 f([\boldsymbol{MeOH}] \cdot [\boldsymbol{X_0}])$	0.823	0.779

$$f([\boldsymbol{X_0}]) = \frac{[X_0] - 112.5}{22.5} \qquad f([\boldsymbol{MeOH}]) = [MeOH] - 1.5 \qquad f(\boldsymbol{t}) = \frac{t - 31}{15}$$

Based on the data generated for the second screening for the Mut+ phenotype, several deductions can be made as a result of changing screening ranges for the investigated process parameters.

Comparing the data sets generated for the first and second screening (Table 10-3 and 10-5 respectively), a significant contrast becomes apparent. The chosen ranges in the first screening resulted in predominantly high expression, yet negative growth rates for all observations. The second screening shows the exact opposite result. The observed expression is so low that it is indistinguishable from noise, reflected by the lack of significant model terms shown in Table 5-9 while for all but one run show positive growth rates[12]. It therefore seems that, for the Mut$^+$ phenotype within the chosen design space, there is an inverse relationship between growth rate and expression.

[12] This outlier can be explained by the requirement of slightly longer adaptation times to higher –or even toxic- concentrations of methanol (comparing pattern '+--'with '+-+').

This fact that the majority of all variation in growth rate response (Table 5-7) is due to variation in methanol concentration indicates that, at higher methanol concentrations, growth -rather than heterologous protein expression- is preferential for the Mut^+ phenotype. As suggested by other research (Chiruvolu, Cregg and Meagher, 1997; Krainer *et al.*, 2012), this could be explained by the elevated requirement of alcohol oxidase at increased levels of methanol. This creates a scenario in which more resources are allocated to the synthesis of alcohol oxidase rather than the heterologous protein of interest. It is therefore crucial for the remainder of this research to bear in mind that increased induction, counterintuitively, does not necessarily result in higher levels of heterologous production.

In addition to the nutrient competition-effect, described above, another possible explanation for the observation that low levels of recombinant protein are associated with high growth and vice versa, is the consumption of heterologous protein as a carbon source. Therefore, at limiting levels of carbon source (i.e. methanol), *P. pastoris* would consume any heterologous protein already expressed, whereas at high levels of methanol, the required overexpression of AOX and increased growth rates preventing toxicity, could result in limited resources being redirected to the synthesis of AOX rather than heterologous protein. The result of these two scenarios is that there could be a tight window of operation for methanol feeding, one that would be particularly difficult to control with bolus methanol feeding.

In summary, to address the hypotheses stated at the start of this section regarding the Mut^+ phenotype: Extending the incubation time has likely led to *P. pastoris* Mut^+ to adapt to higher specific levels of methanol, i.e. the amount of methanol available per unit of biomass. However, this adaptation, unexpectedly, did not lead to higher heterologous protein expression, but rather higher growth rates. In addition, methanol toxicity was shown to occur at low cell concentrations, high methanol concentrations and at relatively early induction times.

Extending the incubation time, increasing the concentration of initial biomass and widening the range of methanol concentration has resulted in significantly higher levels of both expression and cell growth for the Mut^S phenotype.

This illustrates an important difference between Mut^+ and Mut^S phenotypes. For Mut^S phenotypes the correlation between specific growth rate and heterologous protein expression has a propensity to be positive, whereas for the Mut^+ phenotypes this correlation tends to be inversed.

This difference can be explained through Figure 5.3, which is a simplified version of an illustration of the methanol utilisation pathway by Krainer et al. (2012) and incorporates the explanation by Chiruvolu, Cregg and Meagher (1997) of the previous mentioned nutrient-competetion effect between recombinant protein and AOX proteins. Mut^S is known to express considerably less AOX protein than Mut^+. This means that nutrient resources, such as a limited carbon source, would be available for the production of recombinant protein in Mut^S at equal expression

stimuli. This could explain why, in all the screenings performed, volumetric expression levels of tHBc are highest in the Mut^S phenotype.

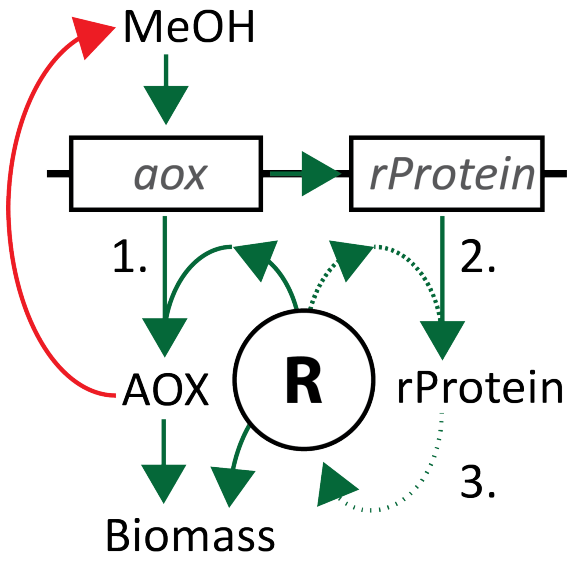

Figure 5.3 Diagram of the methanol utilisation pathway in *P. pastoris* demonstrating nutrient competetion affecting heterologous protein expression. Green arrows indicate positive stimuli; the red arrow shows a negative stimulus. 'R' represents the overall nutrient supply available to a cell. (1) Expression of AOX protein(s) through aox gene(s). (2) Expression of heterolgous protein, dependent on nutrient availibility and induction stimulus. (3) Degradation of heterologous protein.

This section has demonstrated how microtitre plates were used to screen induction stage fermentation for two phenotypes of *P. pastoris*. Key differences between induction behaviour of the two strains were observed and it was demonstrated that, within the chosen design space, the Mut^S phenotype produced the highest volumetric yields of tHBc (see tables 10.3-10.6).

With regards, to the suitability of microtitre plates used to study induction processes, three central limitations were identified for both strains of *P. pastoris*: (1) the requirement for an adaptation period to methanol, (2) the substrate depletion of methanol and (3) methanol toxicity. To address these limitations, either a continuous feeding mechanism needed to be implemented or a bolus

feeding optimisation strategy needed to be developed. The former system was not available in the microtitre plate system and the latter approach would require the addition of significantly more factors added to screening (and eventually optimisation) experimental designs, such as the interval length between the bolus additions and the total length of induction (or the total amount of bolus additions given). Because *P. pastoris* cultures are shown to require an adaptation period to methanol, it would most likely be necessary to implement separate screening experiments for phases where the cultures are not adapted to methanol and when the cultures are adapted to methanol. Also, because the screening experiments shown in this chapter so far have not eliminated a single studied factor as being insignificant, the previous screened factors in addition to the proposed bolus-associated factors would amount to a high number of runs[13] that could be considered uneconomical from a time-management point of view.

5.5 Investigating alternative feeding with pectin and pectin derivatives

The previous section demonstrated the limitations that induction with pure methanol imposes on method development for microtitre plate-based upstream process platforms for *P. pastoris*.

This section aims to address these issues with a question that seems to remain largely unanswered: how does *P. pastoris* assimilate methanol in its natural environment in the first place?

[13] (25+3)·2·2=140 runs for screening experiments, assuming 2-level screening designs with 3 centre points for two phases of induction (adapted and non-adapted) for both phenotypes of *Pichia pastoris*.

P. pastoris was first isolated from a French chestnut tree and was often found as a contaminating yeast species in wine production (BioGrammatics, 2013). This implies that these microorganisms obtain nutrients from plant derivatives. In fact, many methylotrophic yeasts obtain methanol indirectly from a plant polymer: pectin. This happens through the conversion of pectin by pectin methyl esterase (PME) to carboxylated polymers and free methanol (Lee, Jae-Dong and Komagata, 1980; Nakagawa *et al.*, 2005), as shown in Figure 5.4.

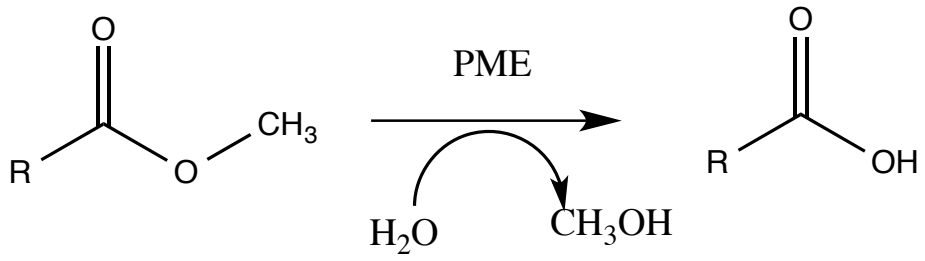

Figure 5.4 Reaction diagram of the enzymatic conversion of a methyl ester on pectin to a carboxylated polymer and free methanol by pectin methyl esterase.

Previous research has also suggested that esters are converted to methanol in *P. pastoris* through an unidentified pathway (Krainer *et al.*, 2012). Taking this and the phylogenetic similarity to other methylotrophic yeasts into account, it was hypothesised that *P. pastoris* may be able to metabolise pectin to obtain methanol.

If this hypothesis were proven to be true, pectin, or its derivatives, could be used as an indirect induction agent. This would mitigate the methanol feeding-associated limitations of microtitre plate-based induction of *P. pastoris* previously discussed in section 5.4.2 and the methanol toxicity issues in unintended accumulation events.

5.5.1 Micro fermentations of pectin digest

It should be stressed that in this series of experiments methanol is still the inducing agent and that pectin would effectively act as molecular storage of methanol. Therefore, it had to be determined how much pectin would be appropriate to produce a significant induction response.

Pectin is polymer of D-galacturonic acid and D-galacturonic acid methyl ester (Mohnen, 2008). The level of esterification, denoted in %, indicates the proportion of molecules in a pectin polymer that are D-galacturonic acid methyl ester. D-galacturonic acid methyl ester (dGAME), specifically, has been postulated to be the methanol donating group within pectin.

It was shown in Figure 5.2 that a minimum observed concentration of methanol required to sustain a culture was 0.313%. Therefore, an equivalent amount of 0.5 % vol. methanol was prepared with pectin from apple. This equivalent was found to translate to 3.67g/L pectin, assuming a 70% level of esterification.

However, pectin is poorly soluble in water and turns to gels at higher concentrations (data not shown). Increased viscosity would provide unfavourable culturing conditions, as this would lower oxygen transfer conditions. An approach to solubilise pectin was to heat pectin under alkaline conditions and subsequently depolymerise pectin using pectin lyase (PNL) as shown in Figure 5.5.

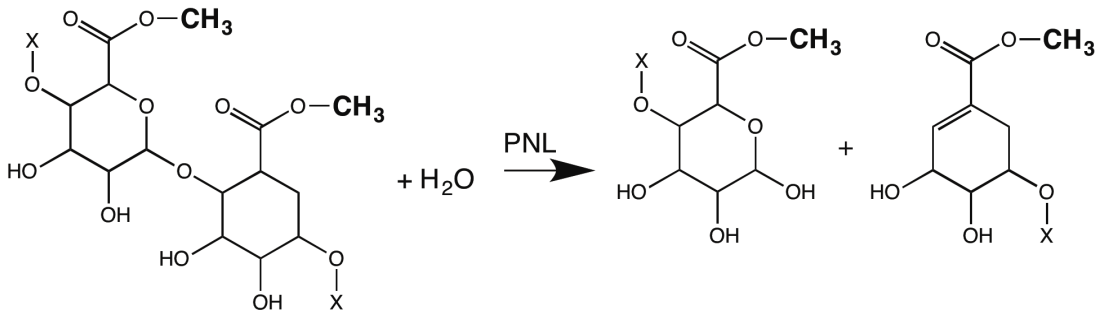

Figure 5.5 PNL-mediated depolymerisation of pection (left) to dGAME (right).

To study if pectin could induce *P. pastoris* indirectly, Pectin Digest Medium (PDM; 0.5% vol. MeOH equivalent dGAME) was prepared as described in 2.3.1.1.

Harvest times (*t*) and initial biomass concentrations ($[X_0]$) were varied in a similar screening-type fashion as described in previous section for both the Mut^+ phenotype (Table 10-7) and the Mut^S phenotype (Table 10-8).

Table 10-9 and Table 10-10 show the level of overall response variance created by each factor in screening experiments for Mut^+ and Mut^S phenotypes respectively. The left side of Table 10-11 and Table 10-12 - Mut^+ and Mut^S phenotypes respectively- show the goodness of fit of these models, using r^2. The robustness of these models was subsequently assessed using r^2 *adjusted*. The same tables subsequently show formulae for each response model showing only significant model terms ($p<0.05$). Similarly, r^2 and r^2 *adjusted* of these simplified models are shown on the right-hand side of these tables.

Mut^+ showed positive growth rates after each experimental run, with low expression of heterologous protein. However, Mut^S showed slight negative growth rates, with significant expression of heterologous protein. This contrast between

the two phenotypes is seen in previous sections and can also be explained by the previously discussed nutrient competition-effect depicted in Figure 5.3.

Overall the use of PDM shows to induce *P. pastoris* proving the hypothesis that pectin derivatives can be used to induce heterologous protein expression.

However, PDM is an undefined enzymatic digest. Previous work has shown that the saponification of citrus pectin through alkalisation could result in complete de-esterification (Zhan, Janssen and Mort, 1998). It was therefore not definitive if the observed effects were due to the action of methanol liberation by PME on dGAME polymers.

5.5.2 Micro fermentations of pectin monomers

The previous section described how Pectin Digest Media was used successfully to induce heterologous protein expression in two phenotypes of *P. pastoris*.

This section describes using defined media conditions to test the hypothesis of monomeric dGAME being able to induce heterologous protein expression through PME activity. To explain the effect of other pectin digest components and to provide a benchmark of expression, d-galacturonic acid (dGA) and methanol were also considered.

Molecularly equivalent amounts to 1% MeOH (vol.) of dGA and dGAME were prepared in BSM. Several mixtures were prepared in which 2-level combinations of the three substrates were combined. To gain an indication of specific substrate

consumption rates, different levels of initial biomass were used for induction. The induction was carried out as described in section 2.3.1.

Figure 5.6 shows the levels of expression and substrate consumption for each possible condition. Expression was measured using dot blotting (2.5.10) followed by immunostaining and linear signal analysis (2.5.11). Residual dGA was measured using uronic acid analysis as described in section 2.5.15, and residual dGAME and methanol were quantified using chromatographic methods described in section 2.5.14.

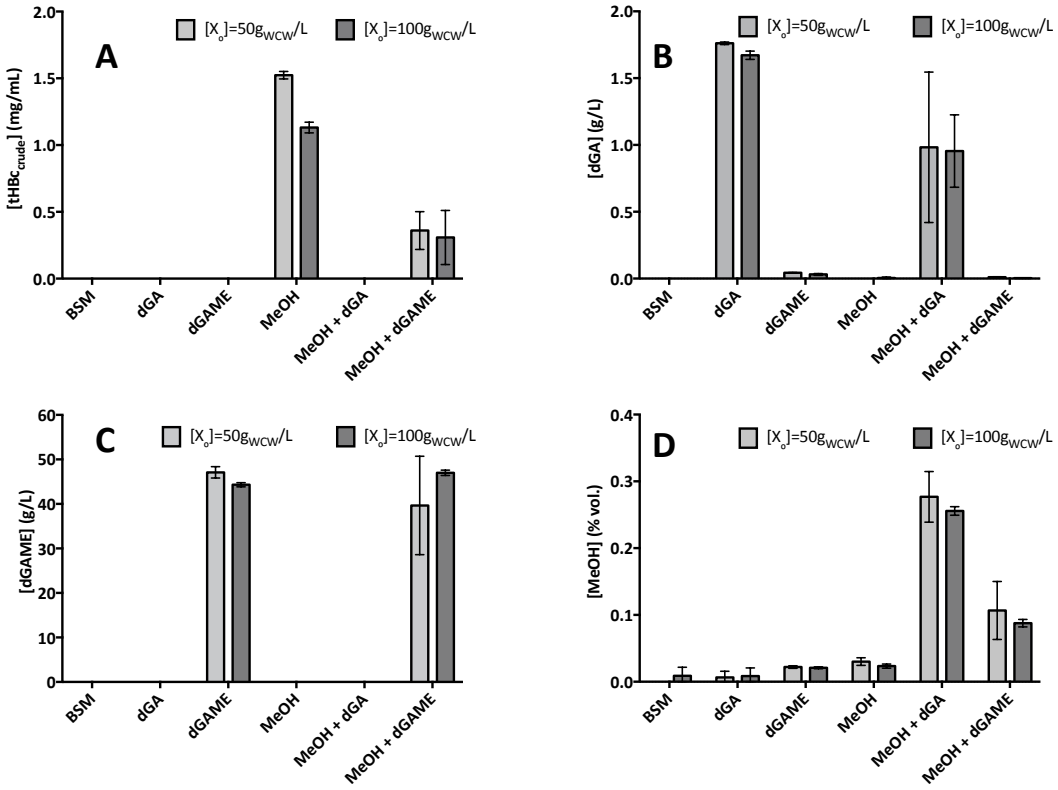

Figure 5.6 Metabolic analysis of *P. pastoris* Mut^S in several defined media conditions: Basal Salts Media (BSM), d-galacturonic acid in Basal Salts Media (dGA), d-galacturonic acid methyl ester in Basal Salts Media (dGAME), methanol in Basal Salts Media (MeOH), d-galacturonic acid and methanol in Basal Salts Media (MeOH + dGA), and d-galacturonic acid methyl ester and methanol in Basal Salts Media (MeOH +dGAME). Expression and susbtrate concentrations after 20 hours of fermentation at two different starting biomass concentrations: (A) Expression of tHBc-k1,k1; (B) dGA concentrations; (C) dGAME concentrations; and (D) MeOH concentrations.

The data shows that induction was only observed in cultures with methanol addition (Figure 5.6A). In all cultures in which methanol was introduced, significant levels of methanol were consumed (Figure 5.6B). No apparent induction was observed in cultures containing dGAME alone (Figure 5.6A) nor was dGAME consumed (Figure 5.6C), disproving the hypothesis that, at least under these conditions, dGAME can be used to induce AOX.

In fact, dGA and dGAME seem to inhibit the uptake of methanol, as shown in Figure 5.6D. This could be explained by the effect that under acidic conditions[14] an equilibrium exists between dGAME on one end; and dGA and methanol on the other end Figure 5.7. This could allow dGA (and dGAME) to act as a 'buffer' for methanol.

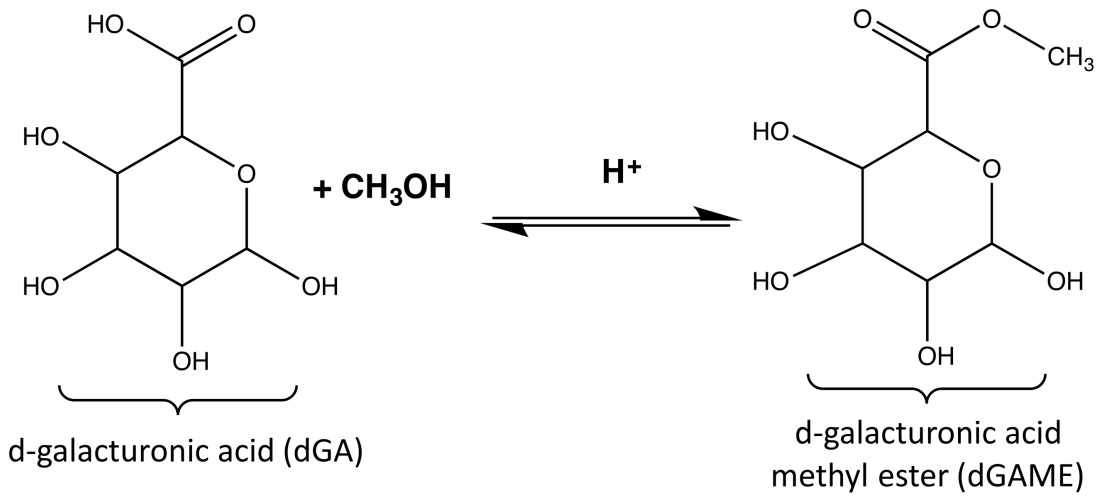

Figure 5.7 Equilibirum between d-galacturonic acid and d-galacturonic acid methyl ester. A reaction mechanism of this equilibrium is shown in Figure 10.10 in the appendix.

[14] Late stage induction cultures in microtitre plates would be significantly below pH 5 due to the consumption of ammonia.

One reason induction was not observed using defined pectin derivatives could be that dGAME is too small a molecule for PMEs to act on to release methanol. Indeed, a review suggest that PMEs exclusively act on chains of galacturonic acid and galacturonic acid methyl ester (Manmohit Kalia, 2015). Therefore, it is hypothesised that galacturonic acid methyl ester can only be used as an indirect inducing agent in polymer, or at least oligomeric form. However, further work would be required to demonstrate this.

5.6 Summary

This chapter studied the use of microtitre plates to develop a miniaturised fermentation platform intended for rapid upstream process characterisation.

It was found that microtitre plates can be used to grow *P. pastoris* for the purposes of increasing biomass in complex media. However, induction of heterologous protein in microtitre plates proved to be more challenging due to the requirement for an adaptation period to methanol, the substrate depletion of methanol and methanol toxicity.

Nonetheless, several critical process parameters were confirmed: Initial biomass concentration at induction, harvest time and methanol concentration in media.

Also, significant differences in growth and induction characteristics between Mut^S and Mut^+ phenotypes were identified. For the Mut^S phenotype the correlation between specific growth rate and heterologous protein expression tended to be

positive, whereas for the Mut^+ phenotype this correlation was inversed. It was postulated this was due to an intracellular nutrient competition effect between expression of AOX protein and heterologous protein.

To address the issues associated with pure methanol-based induction in microtitre plates, the use of pectin digest as an indirect induction agent was studied. It was found that the use of this novel media resulted in the expression of heterologous protein, however, it was determined that this effect was not due to the proposed metabolism of d-galacturonic methyl ester by pectin methyl esterase. Investigating the mechanism of induction by PDM needs to be assessed in future work but could likely be attributed to the release of free methanol during the saponification process of pectin or due to pectin methyl esterases acting exclusively on polymeric forms of d-galacturonic methyl ester.

While microtitre plates have proved useful in previous published work to be an effective screening tool for upstream processing, in the case of methylotrophic *P. pastoris* they have a limited utility. This is because the methanol induction system needs to be tightly regulated and controlled, which cannot be achieved through bolus additions. Therefore, the next chapter will explore miniature bioreactor systems that are capable of tighter parameter regulation.

Chapter 6 Characterisation of Tandem Core Virus-Like particle production

6.1 Introduction

Several critical process parameters were identified in chapter 3 and 5, including temperature; harvest time; initial biomass concentration; and induction strategies.

Chapter 4 shows how microscale sample processing techniques were developed to enable high-throughput sample processing in upstream process development at miniaturised scales. Subsequently, chapter 5 aimed to investigate the use of microtiter plates for the purpose of upstream process characterisation. However, it was found that (1) the oxygen transfer rates required for methanol induction in microtitre plates was limiting and (2) this platform lacked several crucial control mechanisms required for the induction of heterologous protein of *P. pastoris*. Such a platform would therefore be difficult to scale-up to attain practical and industrial relevance.

To address these, this chapter will describe how upstream process characterisation was performed using a novel, miniaturised, disposable and automated bioreactor system, the ambr® 250.

6.2 Aim and objectives

This chapter aims to use miniature bioreactors and the aforementioned scale-down cell disruption tools to study VLP production in a bioprocessing context through variance of previous identified critical process parameters. The specific objectives are:

- Model metabolic effects due to process variance.
- Characterise specific expression of Tandem Core material
- Characterise volumetric expression of Tandem Core material
- Investigate methanol toxicity
- Optimise the VLP production process
- Establish a correlation between VLP assembly and tHBc expression

6.3 Experimental Approach

To better understand VLP expression in *P. pastoris*, the k1,k1 construct investigated in chapter 3 will provide a good base case for the purpose of Tandem Core VLP assembly in this chapter.

This chapter will address the variance of CCPs identified in chapters 3 and 5 through three response categories: metabolic effects, specific expression and volumetric expression.

6.3.1 Metabolic effects studied in the miniature bioreactor system

To gain understanding in the fermentation process, several metabolic responses were studied in the ambr® system, these are: the concentration of biomass [X], the Carbon Evolution Rate (CER), and the Respiratory Quotient (RQ).

6.3.1.1 Concentration of biomass

Using the concentration of biomass to monitor fermentation processes has been covered extensively in previous chapters. This chapter will use wet cell weight (2.5.2) and dry cell weight (2.5.3) to measure biomass in an offline fashion.

6.3.1.2 Carbon Dioxide Evolution Rate

The Carbon Dioxide Evolution Rate (CER) is the net quantity of carbon dioxide observed in the exhaust gas of a fermentation process, expressed as the amount of millimoles of carbon dioxide per litre of working volume per hour. Hence:

$$CER = \frac{F_{in}}{V_W} \cdot \left([CO_2]_{out} \cdot \left({[I]_{in}}/{[I]_{out}} \right) - [CO_2]_{in} \right)$$

Where:

- F_{in} is the sparge rate (m^3/h)
- V_W is the working volume (m^3)
- $[CO_2]_{out}$ is the concentration of exhaust carbon dioxide (mmol/L)
- $[CO_2]_{in}$ is the concentration of sparged carbon dioxide (mmol/L)

- I_{out} is the concentration of exhaust inert gas[15] (mmol/L)
- I_{in} is the concentration of sparged inert gas (mmol/L)

CER analysis is typically used to indicate the metabolic status of a fermentation process through online measurements (Omstead, 1990).

In this series of experiments, CER will be used (1) as a method of monitoring the activity and growth of *P. pastoris*, (2) as a trigger for automatically initiating the glycerol fed-batch phase and (3) comparing performance between scales in the subsequent chapter.

6.3.1.3 Respiratory Quotient (RQ)

The Respiratory Quotient has typically been used as an online means to indicate the carbon source being utilised by a microbe at any given time (Omstead, 1990). The RQ is defined as the ratio between the carbon dioxide produced by a culture and the amount of oxygen consumed by it. Hence:

$$RQ = \frac{[CO_2]}{[O_2]} = \frac{d}{a} = \frac{CER}{OUR}$$

6.3.2 Specific expression

Specific expression is defined as the amount of heterologous protein expressed per unit of biomass. In the case of this study this specifically means the amount of tHBc per gram of wet cell weight: $tHBc_S$ ($\mu g/g_{WCW}$).

[15] Usually argon or molecular nitrogen

One objective of this chapter is to characterise VLP assembly. To do this, the specific expression of tHBc must be characterised in different states. As shown in Figure 6.1, this was done by measuring (1) how much tHBc was recovered after homogenisation, (2) how much of this material was soluble; and (3) how much of this material was in the correct size range of a VLP. These 'states' of tHBc, which are categorised by the downstream fractions in which tHBc is recovered, will henceforth be referred to as 'total' (or 'crude'), 'soluble' and 'VLP' (Figure 6.1)

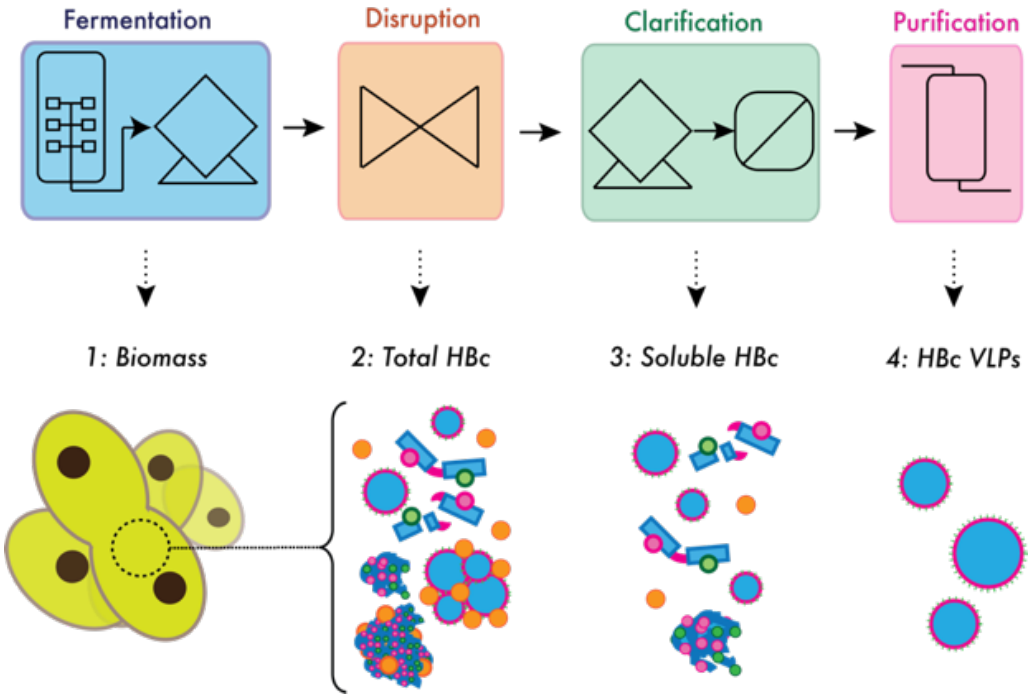

Figure 6.1 Overview of VLP production. (Top) unit operations involved from left to right: Fermentation and harvesting, cell disruption, clarification, and purification. (Bottom) Outputs of respective unit operations: (1) Biomass, (2) lysate containing cell debris, host cell protein and total produced tHBc; (3) clarified lysate containing soluble host cell protein and soluble tHBc; (4) size-exclusion chromatography fractions containing tHBc and host cell protein in within the 20-220µm size range.

Total expression is defined as the quantity of tHBc measured in crude lysate after cell disruption (Step 2, Figure 6.1). Soluble expression is defined as the amount of tHBc measured in clarified lysate after clarification (Step 3, Figure 6.1). Finally, the

amount of tHBc assumed to be mostly assembled VLP was measured after a size exclusion purification step (Step 4, Figure 6.1). It must be emphasised that the chromatography fractions also contained host cell protein, and states of tHBc other than tHBc assembled into Virus-Like Particles. Quantification of tHBc was performed as described in sections 2.5.10 and 2.5.11.

6.3.3 Design of Experiments

This chapter will be addressing the use of more sophisticated tools than microtiter plates for upstream process characterisation. However, to gain higher level of process control, a significant trade-off had to be made for experimental throughput. Bioreactor fermentation processes are considerably more expensive and time consuming than microtiter plate-based fermentation. Therefore, to mitigate the reduced throughput, more economic use of experiments was crucial. Hence, Design of Experiment-based response surface methodology was used for upstream process characterisation.

Figure 6.2 shows an experimental overview of the procedures involved in this chapter to study the above responses.

The effects of temperature, T_i (°C), methanol flow rate, Q_{MeOH} (mL/L_i/h) and finally, glycerol flow rate, Q_{GY} (mL/L_i/h) on responses were studied by generating response surface models. Central Composite Design (CCD) with off-face axial points (α=1.1) and three centre points were used to generate quadratic response surface models. Varying glycerol flow rates allowed for introducing automated variation in biomass

prior to culture induction without relying on offline data. Note that, the centre points of this design space were based on Invitrogen's protocol for *P. pastoris* Mut^S induction (Invitrogen Corporation, 2002).

Models were generated using stepwise regression to maximise the number of model terms while minimising RMSE values.

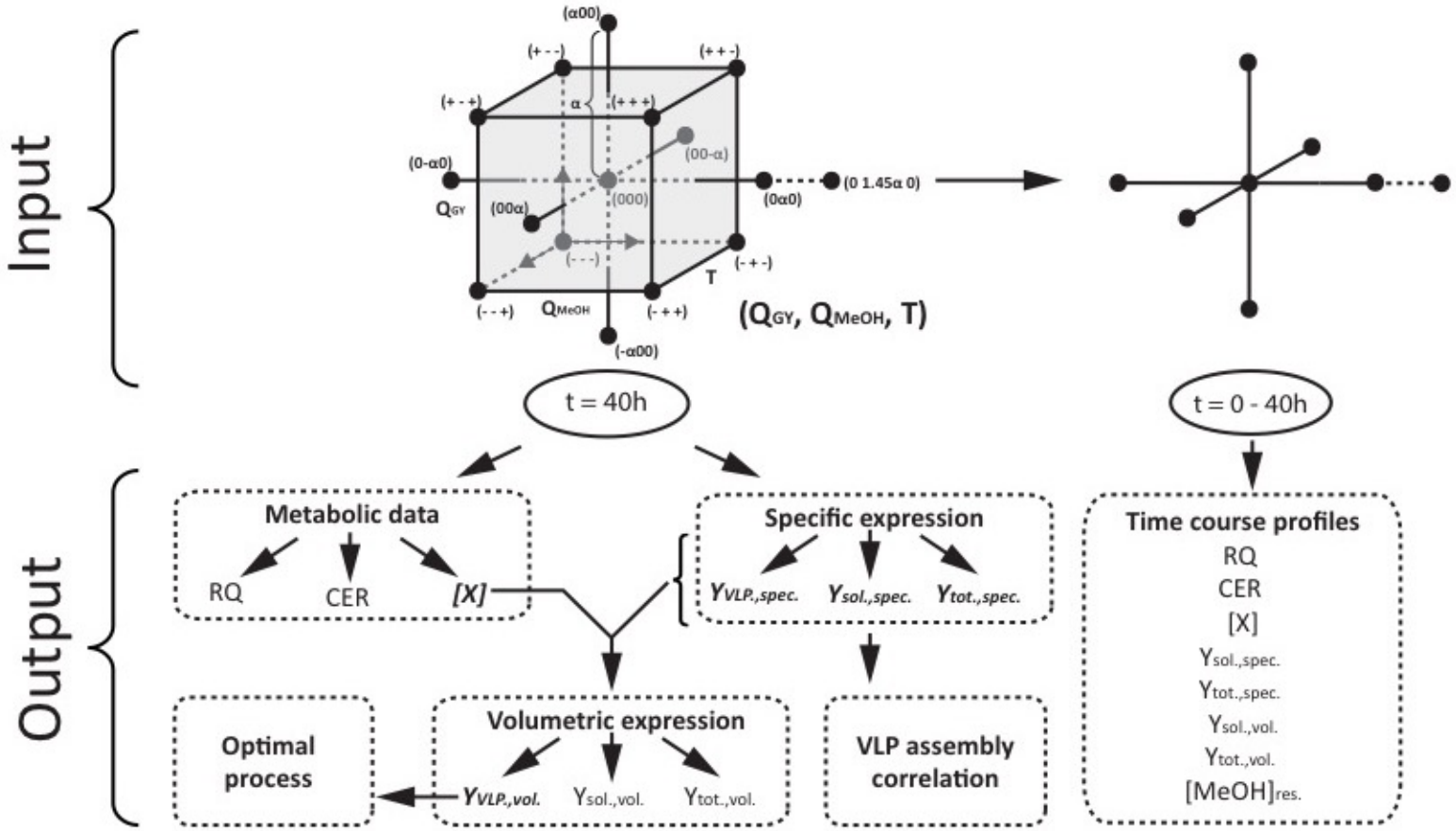

Figure 6.2 Experimental overview. Input: (top left) Central Composite Design for Q_{GY}, Q_{MeOH}, and T after 40 hours of induction; (top right) expansion and centre points of CCD used to investigate response variance over time at 0, 8, 24 and 40 hours of induction. Outputs: Metabolic data including respiratory quotient, Carbon Dioxide Evolution Rate and biomass concentration; and specific expression of various fractions of tHBc.

6.3.4 Methodology

Figure 6.3 describes the workflow used in this series of experiments. Step A described the inoculum culture required to generate biomass for unit operation B, the fermentation unit operation. Variation of process parameters of this unit operation, particularly the induction phase, will be the focus of this chapter. Beyond step B, where the workflow diverges, the top row represents unit the operations involved in developing surface response models; the bottom row represents the unit operations involved in developing time course profiles of each investigated response. The main differences between the top and bottom processing sequences are (1) the use of small sample volumes (40mL and 1.8mL respectively), (2) the method of cell disruption (HPH vs lyticase treatment followed by AFA) and (3) the absence of a size-exclusion chromatography step in the bottom sequence.

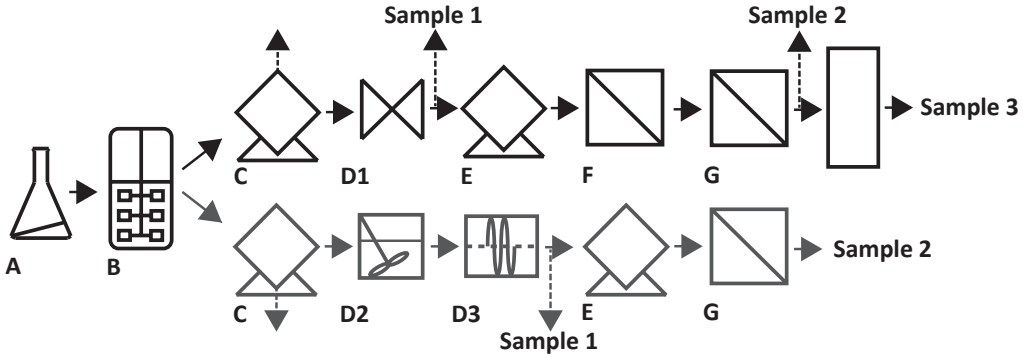

Figure 6.3 Overview of methodology. (A) Shakeflask culture as described in 2.3.2. (B) Fermentation process (2.3.3.1), varied as described by Figure 6.2. (C) Centrifugation (2.4.1), pellet is recovered. (D1) Homogenisation at 3 passes and 500 bar, as described in 2.4.3. (D2) Lyticase incubation as described in 2.4.6. (D3) AFA as described in 2.4.5, with Chromacol tubes. (E) Centrifugation (2.4.10), supernatant is recovered. (F) 0.45µm dead-end filtration. (G) 0.22µm dead-end filtration. Sample 1 represents total expression of tHBc. Sample 2 represents expression of soluble tHBc. Sample 3 represents the fraction containing VLP-assembled tHBc.

The latter is because the sample volume did not permit preparative size-exclusion chromatography. Table 10-13 describes the design space used in the chapter to study nine responses. The models for these responses are summarised in Table 6-1.

Table 6-1 Models and goodness of fit for final biomass concentration ([X]), Carbon Dioxide Evolution Rate (CER), Respiratory Quotient (RQ), specific total expression of core protein ($Y_{tot.,\ spec.}$), specific expression of soluble core protein ($Y_{sol.,\ spec.}$), specific expression of VLP-sized core protein ($Y_{VLP,\ spec.}$), volumetric total expression of core protein ($Y_{tot.,\ vol.}$), volumetric expression of soluble core protein ($Y_{sol.,\ vol.}$) and volumetric expression of VLP-sized core protein ($Y_{VLP,\ vol.}$). Goodness of fit is given on the left-hand side of the table ('All terms') using r^2 and adjusted r^2 for models in which all possible first-degree and second-degree polynomial terms have been considered. Following this, are the formulae describing first- and second-degree polynomials containing a reduced number model terms at the lowest model RMSE. Highlighted in bold are internal functions.

	All terms (Complex model)			Significant terms (Simplified model)	
	R^2	R^2 adj.	Formulae	R^2	R^2 adj.
[X] (gWCW/L)	0.871	0.705	$[X] = 235 - 29.1f(\boldsymbol{Q_{MeOH}} \cdot \boldsymbol{T}) - 25.3f(\boldsymbol{T}) - 23.3f(\boldsymbol{T^2}) + 17.4\,f(\boldsymbol{Q_{MeOH}}) + 11.4f(\boldsymbol{Q_{GY}})$	0.866	0.786
CER (mmoL/L/h)	0.964	0.917	$CER = 46.8 + 20.8f(\boldsymbol{Q_{MeOH}}) - 10.6f(\boldsymbol{Q_{MeOH}} \cdot \boldsymbol{T}) - 8.08(\boldsymbol{T}) - 6.75f(\boldsymbol{T^2}) - 3.56f(\boldsymbol{Q_{MeOH}}^2) + 3.00f(\boldsymbol{Q_{GY}} \cdot \boldsymbol{T})$	0.962	0.940
RQ	0.796	0.533	$RQ = 0.554 + 0.0362f(\boldsymbol{Q_{GY}}) + 0.0300f(\boldsymbol{Q_{MeOH}}^2) + 0.0278f(\boldsymbol{T}) + 0.0242f(\boldsymbol{T^2}) - 0.0166f(\boldsymbol{Q_{MeOH}})$	0.743	0.626
$Y_{tot.,\ spec.}$ (µg/gWCW)	0.596	0.0771	$Y_{tot.,spec.} = 4.17 - 1.36f(\boldsymbol{Q_{GY}}^2) - 1.34f(\boldsymbol{Q_{MeOH}}^2) + 0.975f(\boldsymbol{T}) + 0.872f(\boldsymbol{Q_{MeOH}}))$	0.558	0.411
$Y_{sol.,\ spec.}$ (µg/gWCW)	0.590	0.0632	$Y_{sol.,spec.} = 5.41 - 2.02f(\boldsymbol{Q_{MeOH}}^2) - 1.78f(\boldsymbol{Q_{GY}}^2) + 1.12f(\boldsymbol{Q_{MeOH}}) + 1.05f(\boldsymbol{T})$	0.554	0.405
$Y_{VLP,\ spec.}$ (µg/gWCW)	0.558	-0.00945	$Y_{VLP,spec.} = 4.03 - 1.25f(\boldsymbol{Q_{MeOH}}^2) + 1.08f(\boldsymbol{T}) + 1.02f(\boldsymbol{Q_{MeOH}}) - 0.922f(\boldsymbol{T^2})$	0.476	0.302
$Y_{tot.,\ vol.}$ (µg/L)	0.561	-0.00307	$Y_{tot.,vol.} = 980 - 304f(\boldsymbol{Q_{GY}}^2) + 228f(\boldsymbol{Q_{MeOH}}) - 223f(\boldsymbol{Q_{MeOH}}^2) - 167f(\boldsymbol{T^2}) + 137f(\boldsymbol{T})$	0.539	0.329
$Y_{sol.,\ vol.}$ (µg/L)	0.5690	0.0149	$Y_{sol.,vol.} = 1.35 \cdot 10^3 - 383f(\boldsymbol{Q_{GY}}^2) - 344(\boldsymbol{T^2}) - 338(\boldsymbol{Q_{MeOH}}^2) + 290f(\boldsymbol{Q_{MeOH}})$	0.526	0.370
$Y_{VLP,\ vol.}$ (µg/gWCW)	0.508	0.0167	$Y_{VLP,vol.} = 1005 + 273f(\boldsymbol{Q_{MeOH}}) - 359f(\boldsymbol{T^2}) + 153f(\boldsymbol{T}) - 205f(\boldsymbol{Q_{MeOH}}^2)$	0.444	0.259

$$f(\boldsymbol{Q_{MeOH}}) = \frac{(\boldsymbol{Q_{MeOH}} - 3)}{2} \qquad f(Q_{GY}) = \frac{(\boldsymbol{Q_{GY}} - 18.15)}{5} \qquad f(\boldsymbol{T}) = \frac{(\boldsymbol{T} - 25)}{5}$$

6.4 Metabolic responses

Figure 6.6 shows the time course profiles of metabolic responses: biomass, CER and RQ. Each of these responses is discussed individually through additional analysis of surface response methodology in subsequent subsections.

6.4.1 Effect of methanol uptake

To understand the majority of the metabolic responses, the uptake of methanol must be understood. Figure 6.4 shows that there is an uptake limit of residual methanol in the media by the methylotrophic culture and that this limit occurs at an undefined value between Q_{MeOH}=5.5mL/L_i/h and Q_{MeOH}=8.0mL/L_i/h. The Invitrogen *P. pastoris* fermentation protocol recommends not to exceed a 2% residual methanol concentration as this will induce toxicity (Invitrogen Corporation, 2002). At the same time point that this residual concentration is reached, a decline in growth rate (Figure 6.5A), a reduction in CER (Figure 6.5D) and an increase in RQ is observed (Figure 6.5G).

The methanol uptake rate also seems negatively affected by the induction temperature, as shown in Figure 6.4C.

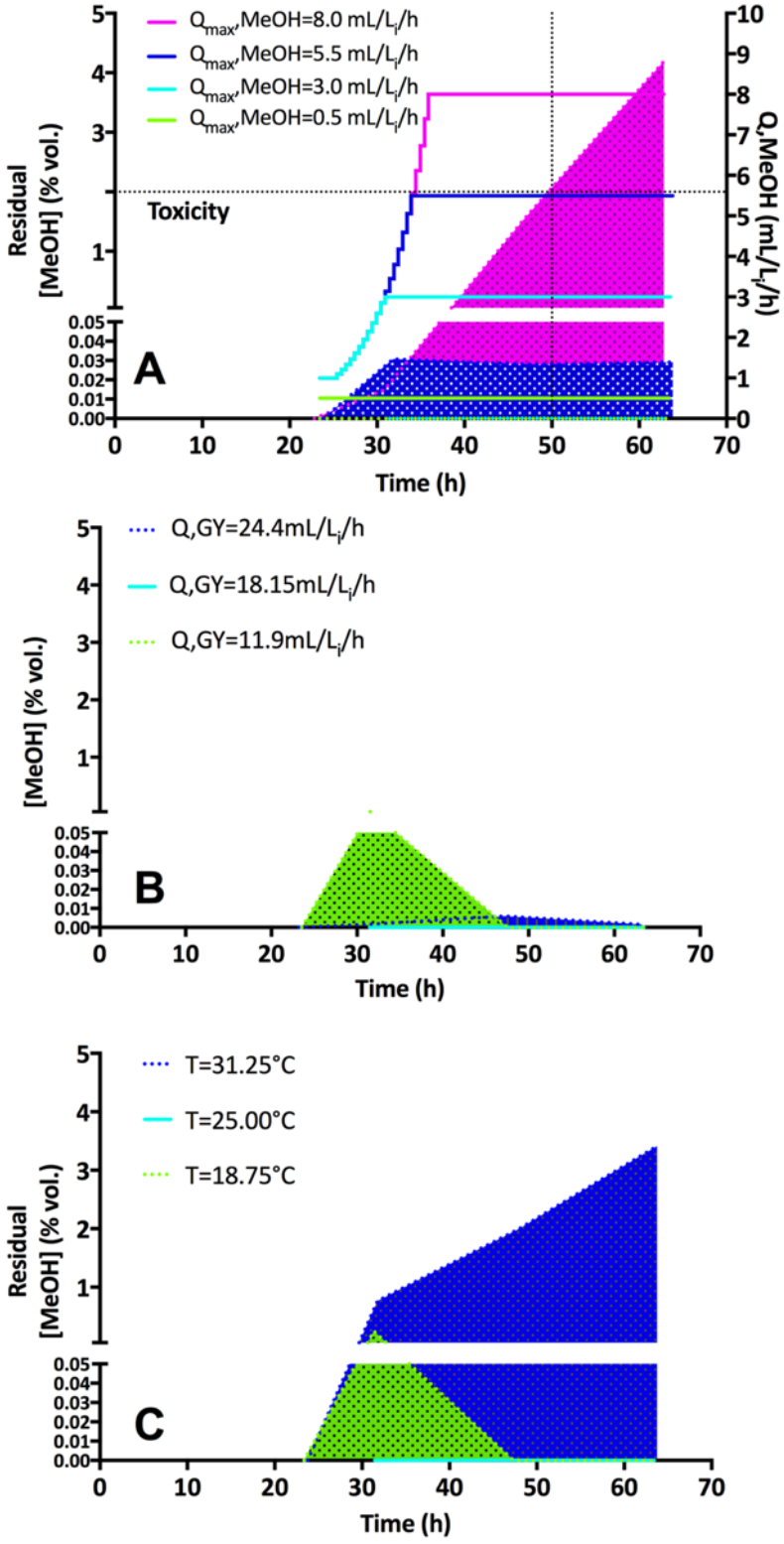

Figure 6.4 Time course profiles of residual methanol concentrations: (A) Residual methanol concentrations at various methanol flow rates. Note that these flow rates, displayed as a line are shown on the right y-axis and the residual accumulation by area plot on the left axis. (B) Residual methanol concentrations at various glycerol flow rates. (C) Residual methanol concentrations at various induction temperatures.

6.4.2 Biomass

Within the ranges of the central composite design, the methanol flow rate seems to have a positive effect on the concentration of biomass (Figure 6.6A&B).

The effect of glycerol flow rate in the fed batch phase seems to have a slight positive effect on the biomass yield after 40 hours of induction (Figure 6.6C&D), however, as indicated by the convergence of growth curves in Figure 6.5B, as induction time is increased the effect of initial biomass concentration, and thus the effect of Q_{GY}, diminishes.

Temperature seems to have a negative effect on overall biomass concentrations (Figure 6.6E&F). This could be explained by the fact that alcohol oxidase activity is increased at higher temperatures (Couderc and Baratti, 1980). This results in the formation of a toxic metabolite, formaldehyde, which in turn could lead to cell death. Other studies have indeed shown that lower induction temperatures lead to a decrease in stress responses, in turn leading to higher biomass concentrations (M Dragosits *et al.*, 2009).

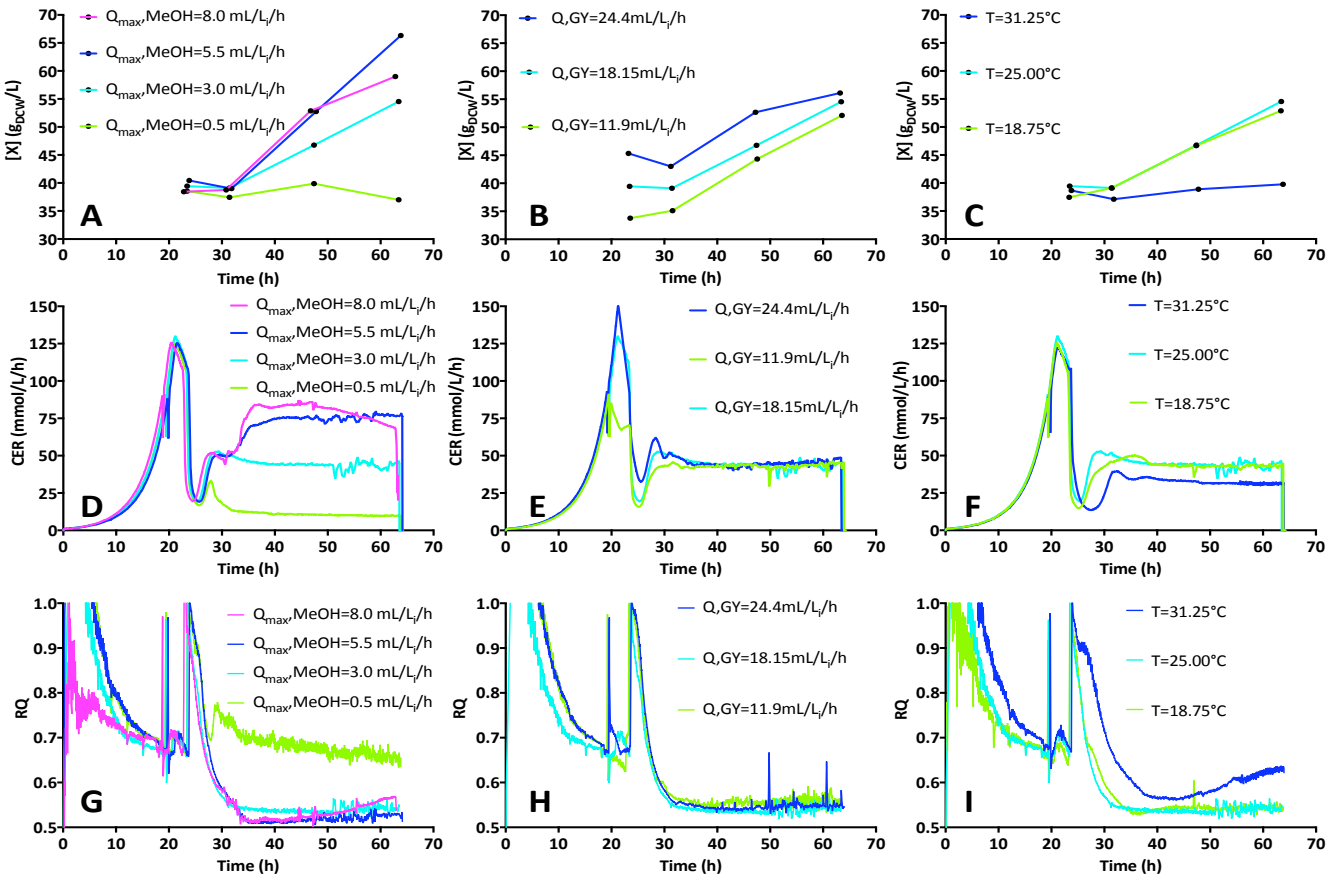

Figure 6.5 Time course profiles of metabolic responses: Biomass (A-C), Carbon Dioxide Evolution Rate (D-F) and Respiratory Quotient (G-I). These responses are a result of varying the following factors: Methanol flow rate (A, D, G); Glycerol flow rate (B, E, H); and induction temperature (C, F, I).

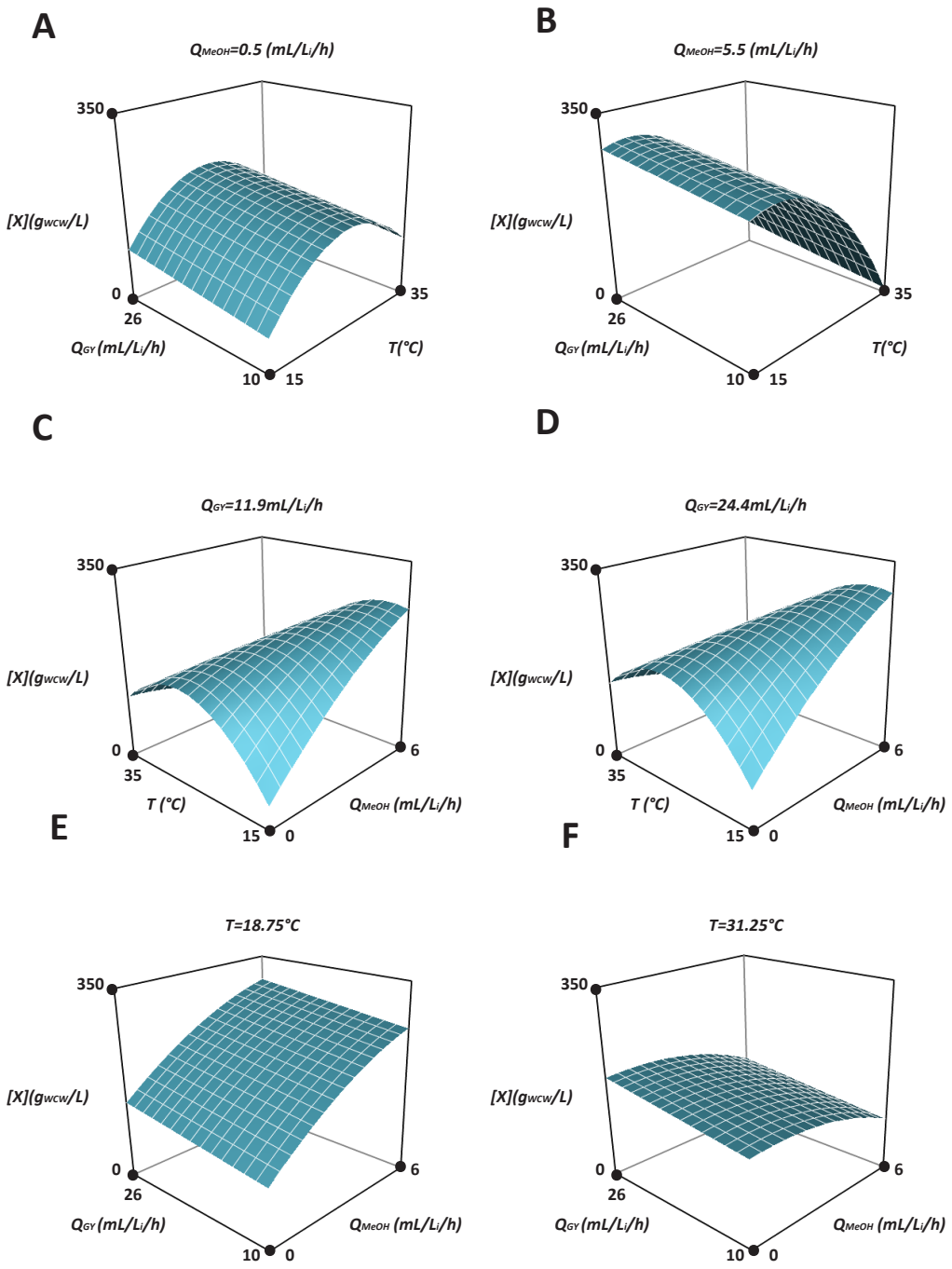

Figure 6.6 Response Surface Model for biomass, [X]. (A-B) Effect of Q_{GY} and T at (A) Q_{MeOH}=0.5mL/L_i/h and (B) Q_{MeOH}=5.5mL/L_i/h. (C-D) Effect of Q_{MeOH} and T at (C) Q_{GY}=11.9mL/L_i/h and (D) Q_{GY}=24.4mL/L_i/h. (E-F) Effect of Q_{GY} and Q_{MeOH} at (E) T=18.75°C and (F) T=31.25°C

6.4.3 The impact of variables on the CER

Methanol feed rate has a positive effect on CER in the investigated design space (Figure 6.7A&B). However, as with biomass when the range of methanol feed rate was extended to promote toxic conditions, a drop in carbon evolution rate was observed in later stages of induction (Figure 6.5D). Indeed, the quadratic term in formula in Table 6-1 indicates that there is an optimum methanol flow rate outside the investigated range beyond the higher end of the flow rate levels.

Glycerol flow rate seems to have a negative effect on CER, interacting with induction temperature.

The induction temperature also has a negative effect on the CER, the lower the temperature, the higher the CER. However, the model indicates that temperature may have an optimum outside the investigated range. However, this value is beyond the lower end of the investigated temperature settings and would impact on product expression.

In addition, note the high goodness of fit of the model in Table 6-1 and the high 'r^2 adjusted' value, indicating that the model has high predicative capacity. This model is therefore useful, not just for the production of recombinant vaccines, but for the performance prediction of any *P. pastoris* culture, assuming negligible effects introduced by varying the heterologous protein of interests (i.e. stress responses, nutrient competition and protein recycling).

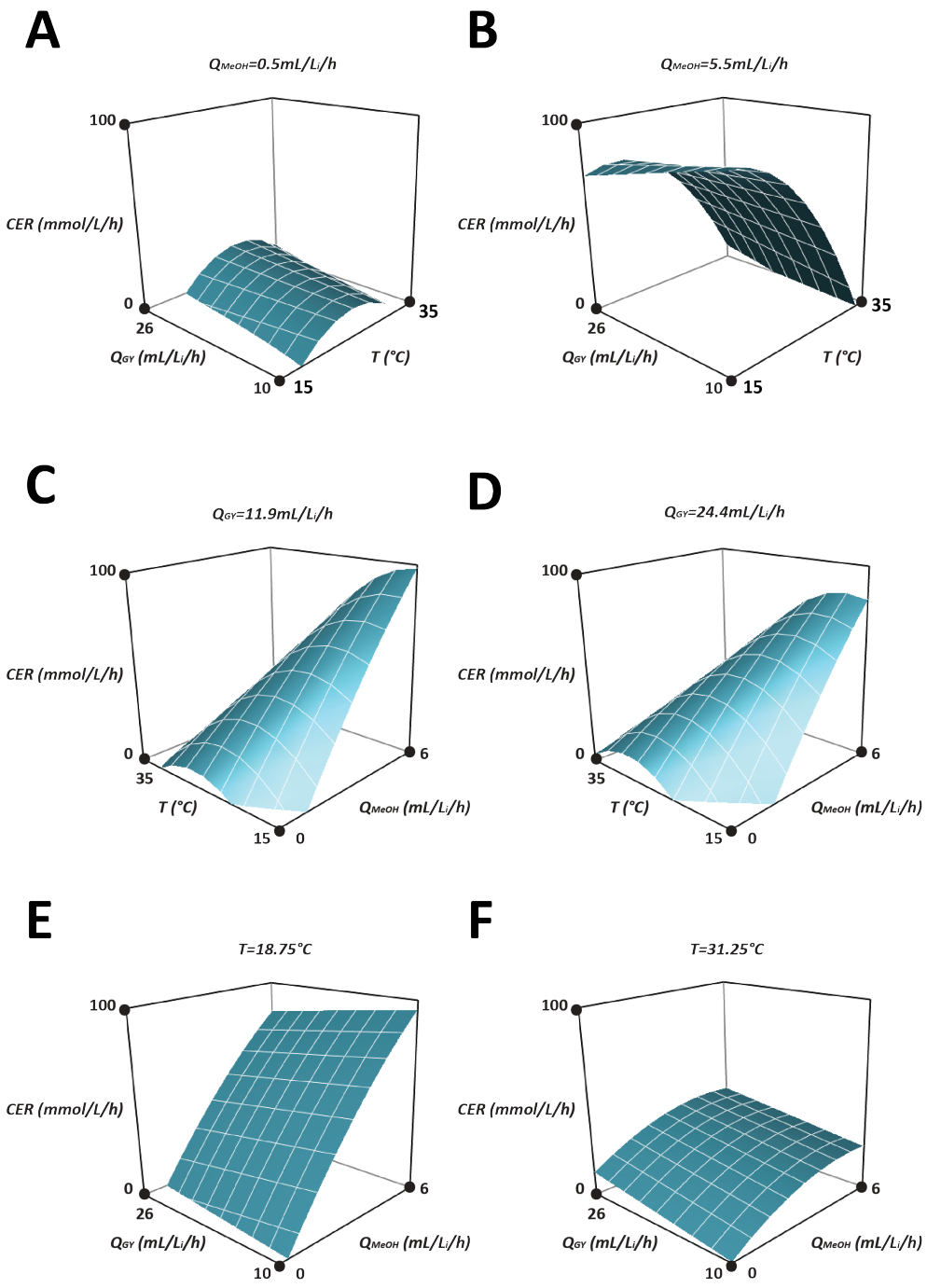

Figure 6.7 Response Surface Model for Carbon Dioxide Evolution Rate, CER. (A-B) Effect of Q_{GY} and T at (A) Q_{MeOH}=0.5mL/L_i/h and (B) Q_{MeOH}=5.5mL/L_i/h. (C-D) Effect of Q_{MeOH} and T at (C) Q_{GY}=11.9mL/L_i/h and (D) Q_{GY}=24.4mL/L_i/h. (E-F) Effect of Q_{GY} and Q_{MeOH} at (E) T=18.75°C and (F) T=31.25°C

6.4.4 Respiratory quotient

The respiratory quotient is significantly influenced by methanol flow rates. As shown in Figure 6.8A-C there seems to be a minimum RQ occurring around 3.55mL/Li/h. Note that the toxic effect of methanol introduced through the 1.45α expansion point has not been incorporated into this model, but is shown in Figure 6.5C. In this last graph, a slight increase in RQ is seen towards the end of the fermentation.

Increases in the glycerol flow rate seem to cause an increase in the RQ in Figure 6.8. However, this effect is not very noticeable in Figure 6.5. Temperature also seems to have a largely positive effect on the value of RQ, with the minimum existing around 22.04°C.

The value of the RQ can be explained as follows. When methanol is only used for maintenance and not being fixed as biomass:

$$2CH_3OH + 3O_2 \rightarrow 4H_2O + 2CO_2$$

Hence: $RQ = {}^{2}/_{3} = 0.67$

This principle is demonstrated in the case of the induction at Q_{MeOH}=0.5mL/L_i/h (see Figure 6.5A & G), where no significant growth is observed and the RQ is indeed around 0.67.

However, when methanol is being fixed as biomass, as is the case during growth, this value can become lower as shown by a simplified version of Equation 10-4, where the assumption is made that carbon fixation via protein expression is negligible[16]:

$$CH_3OH + a \cdot O_2 + b \cdot NH_4OH \rightarrow c \cdot CH_\alpha O_\beta N_\gamma + d \cdot CO_2 + e \cdot H_2O$$

Hence, when $c > 0$, $RQ = {}^{d}/_{a} < 0.67$

[16] This seems like a reasonable assumption as the maximum observed specific expression in this research is around 60µg/gDCW.

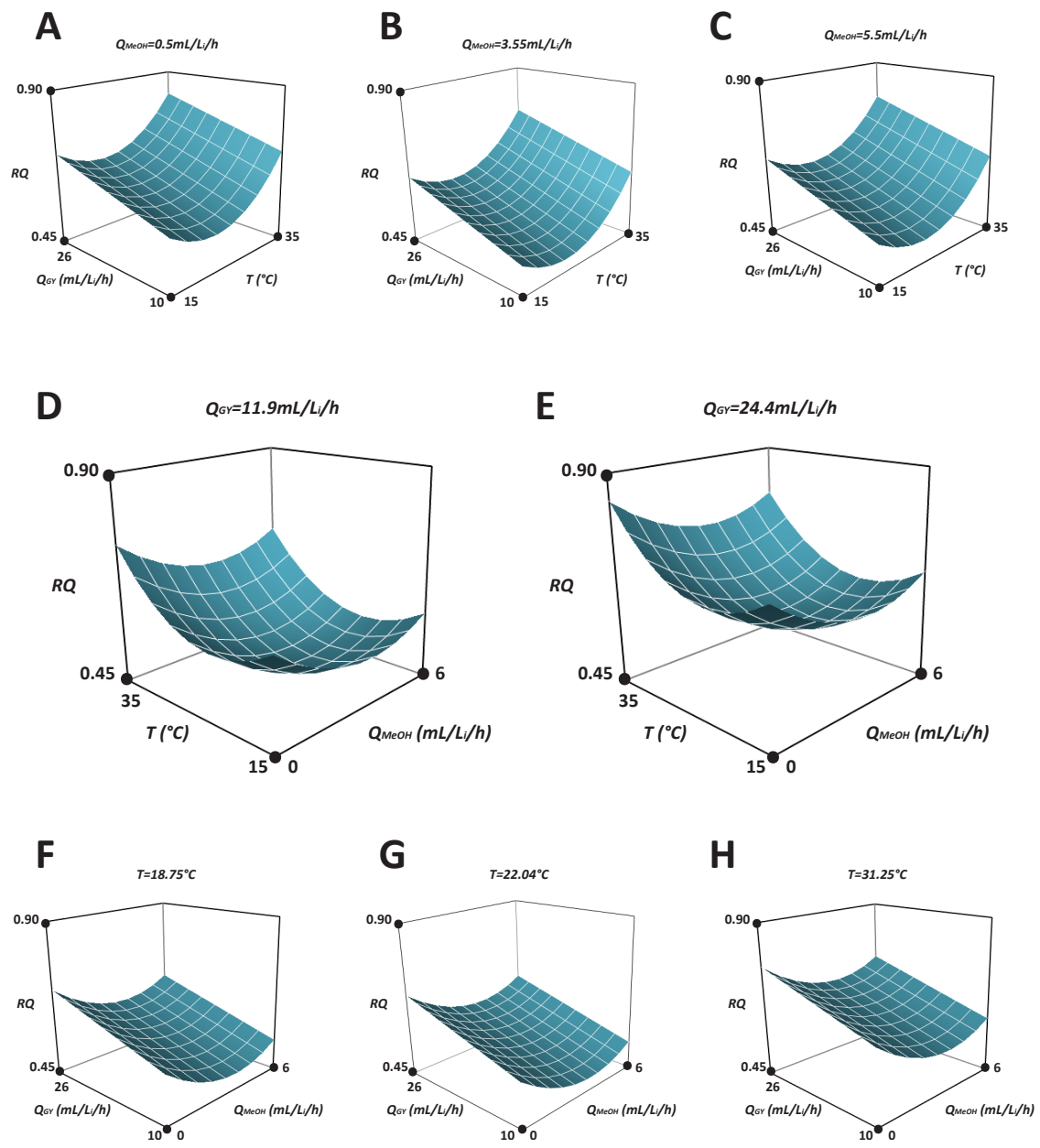

Figure 6.8 Response Surface Model for respiratory quotient, RQ. (A-C) Effect of Q_{GY} and T at (A) Q_{MeOH}=0.5mL/L_i/h, (B) Q_{MeOH}=3.55mL/L_i/h (response minimum) and (C) Q_{MeOH}=5.5mL/L_i/h. (D-E) Effect of Q_{MeOH} and T at (D) Q_{GY}=11.9mL/L_i/h and (E) Q_{GY}=24.4mL/L_i/h. (F-H) Effect of Q_{GY} and Q_{MeOH} at (F) T=18.75°C, (G) T=22.04°C (response minimum) and (H) T=31.25°C

6.5 Specific expression of Tandem Core protein

6.5.1 Specific total core protein

Figure 6.9 shows the response surface model for specific total core protein expression. Optima were established for methanol flow rate (Q_{MeOH}=3.65mL/L_i/h) and glycerol flow rate (Q_{GY}=18.15mL/L_i/h) and a positive correlation was established between temperature and specific expression. The existence of an optimum methanol flow rate for this response can explained by the fact that methanol acts as a carbon source required for growth, yet is toxic at higher concentrations. The optimum glycerol feed rate in the fed batch phase is aliased to the optimum biomass concentration at the start of the induction phase.

The existence of an optimum biomass concentration at the start of the induction phase has a similar explanation as the existence for an optimum methanol flow rate: too low biomass concentration relative to residual methanol concentration could result in cell toxicity, whereas too high concentrations of biomass relative to residual methanol concentration would result in nutrient competition and reduced specific induction stimulus.

The positive correlation of temperature to expression of tHBc could be explained by the previous postulation, made in section 6.4.2, that higher temperatures lead to lower levels of biomass through alcohol oxidase hyperactivity. A decrease in biomass at the same volumetric induction stimulus would lead to higher specific expression.

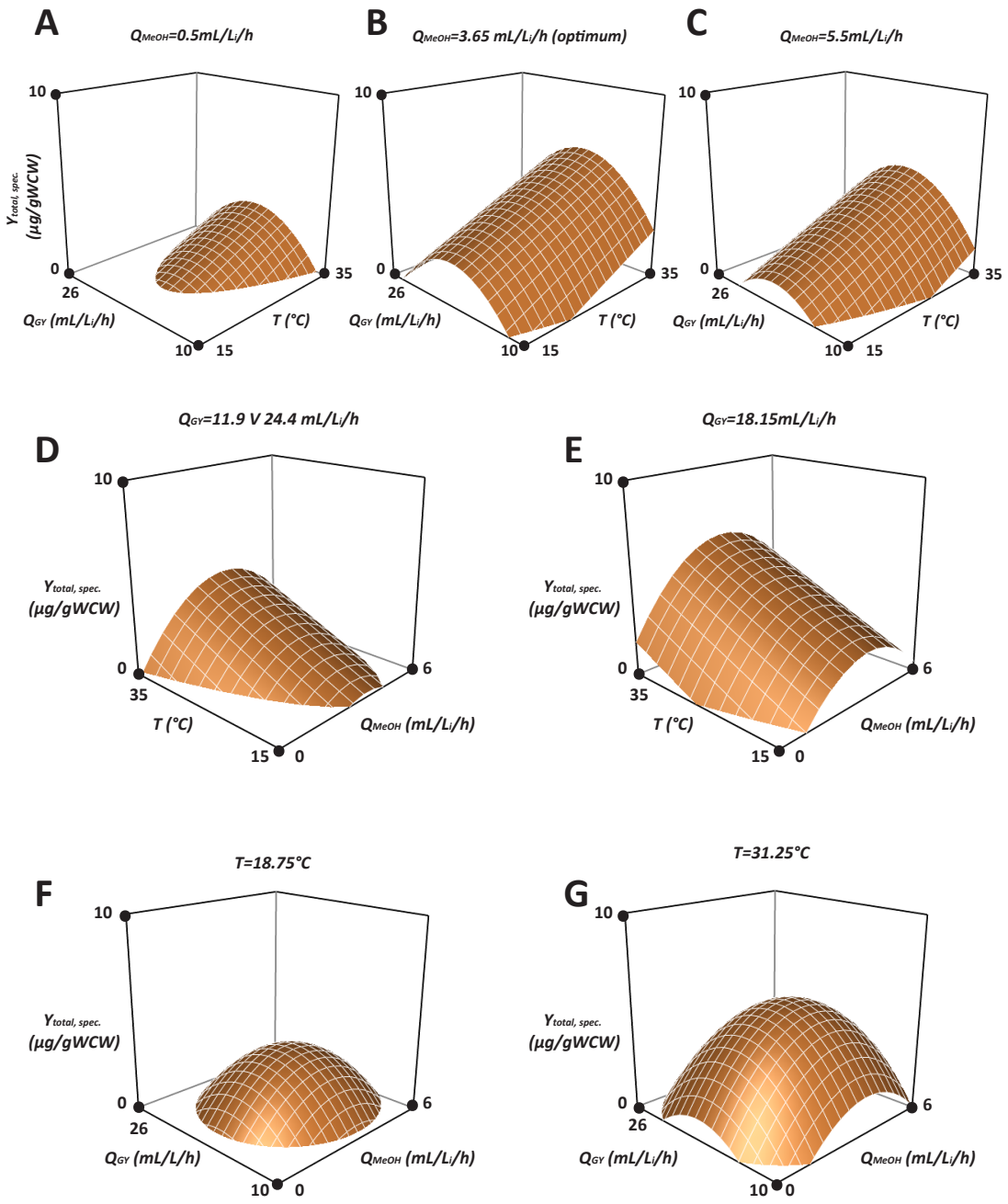

Figure 6.9 Response Surface Model for specific total core protein expression, $Y_{total, spec}$. (A-C) Effect of Q_{GY} and T at (A) Q_{MeOH}=0.5mL/L$_i$/h, (B) Q_{MeOH}=3.65mL/L$_i$/h (response maximum) and (C) Q_{MeOH}=5.5mL/L$_i$/h. (D-E) Effect of Q_{MeOH} and T at (D) Q_{GY}=11.9mL/L$_i$/h and (E) Q_{GY}=24.4mL/L$_i$/h. (F-G) Effect of Q_{GY} and Q_{MeOH} at (F) T=18.75°C, (G) T=31.25°C.

6.5.2 Specific soluble core protein

Specific soluble tHBc is relevant as this is the state of Tandem Core material that can either be immunogenic unassembled or assemble into a VLP.

As seen in Figure 6.10, the trends observed for specific soluble tHBc expression are very similar to that for specific total core protein expression.

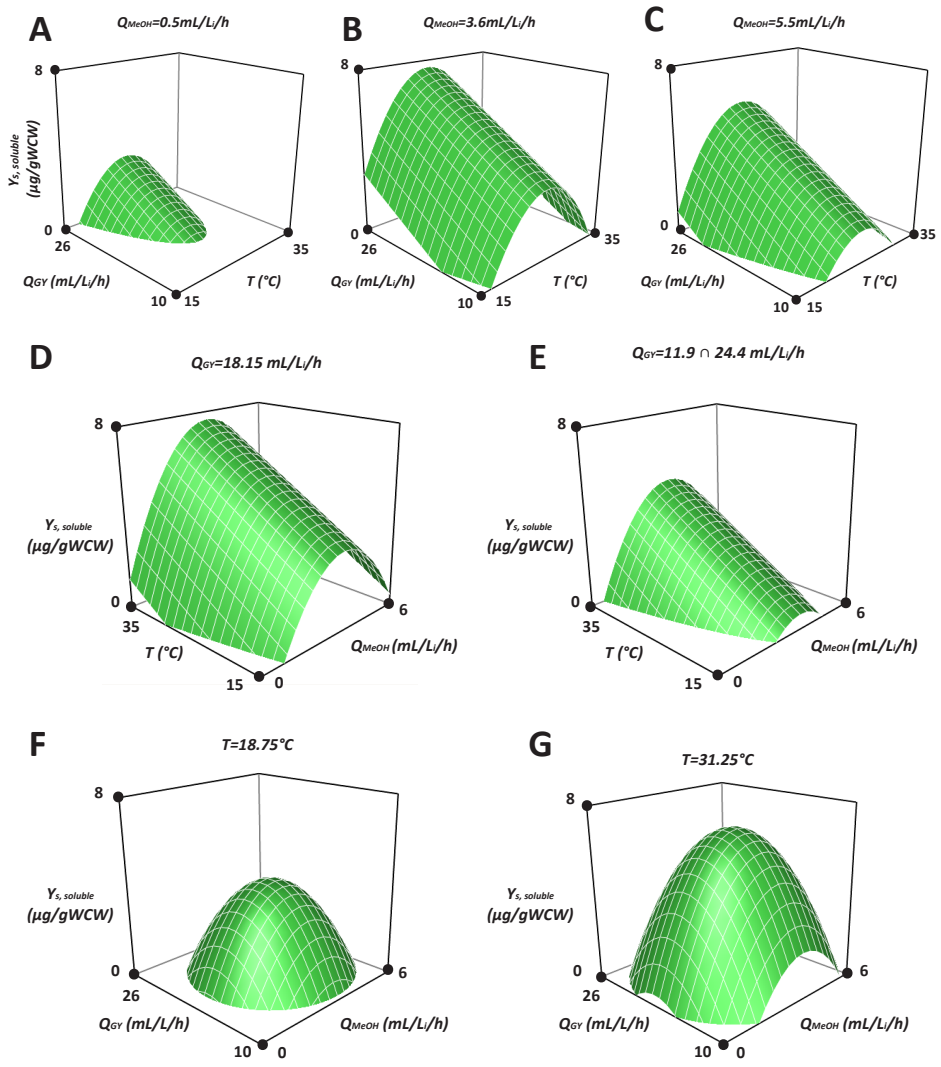

Figure 6.10 Response Surface Model for specific soluble core protein expression, $Y_{sol., spec.}$ (A-C) Effect of Q_{GY} and T at (A) Q_{MeOH}=0.5mL/L$_i$/h, (B) Q_{MeOH}=3.60mL/L$_i$/h (response maximum) and (C) Q_{MeOH}=5.5mL/L$_i$/h. (D-E) Effect of Q_{MeOH} and T at (D) Q_{GY}=11.90mL/L$_i$/h and Q_{GY}=24.40mL/L$_i$/h; and (E) Q_{GY}=18.15mL/L$_i$/h (response maximum). (F-G) Effect of Q_{GY} and Q_{MeOH} at (F) T=18.75°C, (G) T=31.25°C.

6.5.3 Time course profiles of specific expression responses

Figure 6.11 shows time course profiles of total and soluble expression as a result of variation of methanol flow rate, glycerol flow rate and induction temperature. Small-scale sample processing and analysis was performed as described in sections 6.3.4, 2.5.10 and 2.5.11.

Increasing methanol flow rate resulted in higher expression over time of total tHBc and increased amount of soluble tHBc expression. However, this correlation for soluble tHBc changed beyond Q_{MeOH}=5.5mL/L_i/h. At Q_{MeOH}=8.0mL/L_i/h, virtually no production of soluble protein was seen after 24 hours of induction, whilst total expression was the highest observed at this time point. This could suggest that, at very high induction stimulus the resulting expression of tHBc is either incorrect, resulting in truncated heterologous protein, or is too high to permit assembly of soluble multimeric structures. In addition, this flow rate caused observed methanol toxicity at 24 hours after induction. This is indicated by a decreased growth rate (Figure 6.5A), a decreased CER (Figure 6.5D) and an increasing RQ (Figure 6.5G) coinciding with a residual media methanol concentration over 2% vol. (Figure 6.4A), which has been shown by previous work to be the limit toxicity threshold concentration of residual methanol (Invitrogen Corporation, 2002). Methanol toxicity could have triggered various stress responses including the up-regulation of proteases, which would have resulted in truncation of tHBc, contributing to poor solubility or poor identification by dot blot, and poor assembly. In addition, methanol can weaken hydrophobic interactions between protein (Hwang *et al.*,

2011). This in turn can affect the assembly of tHBc into soluble material (such as VLPs) as this action is based on hydrophobic interactions (Holmes *et al.*, 2015; Peyret *et al.*, 2015).

The effect of glycerol flow rate in the fed batch phase, leading to higher biomass concentrations at the start of the induction phase, does not seems to have a significant effect on the solubility of expressed tHBc at later stages of induction.

Temperature seemed to have a positive effect on the total specific expression of tHBc, yet a negative effect on soluble, recoverable heterologous protein expression. It has been reported that the protein folding stress is increased at high induction temperatures and that higher titres of correctly folded protein are often achieved below 30°C (M Dragosits *et al.*, 2009; Li *et al.*, 2011; Anasontzis *et al.*, 2014). The higher temperatures could have therefore led to increased misfolding and hydrophobicity of expressed tHBc, such that, either during the induction process or during the cell disruption process, expressed Tandem Core material associated more readily with host cell protein or cell debris. This would have led to the formation of aggregates entrapping core-positive material, hindering recovery.

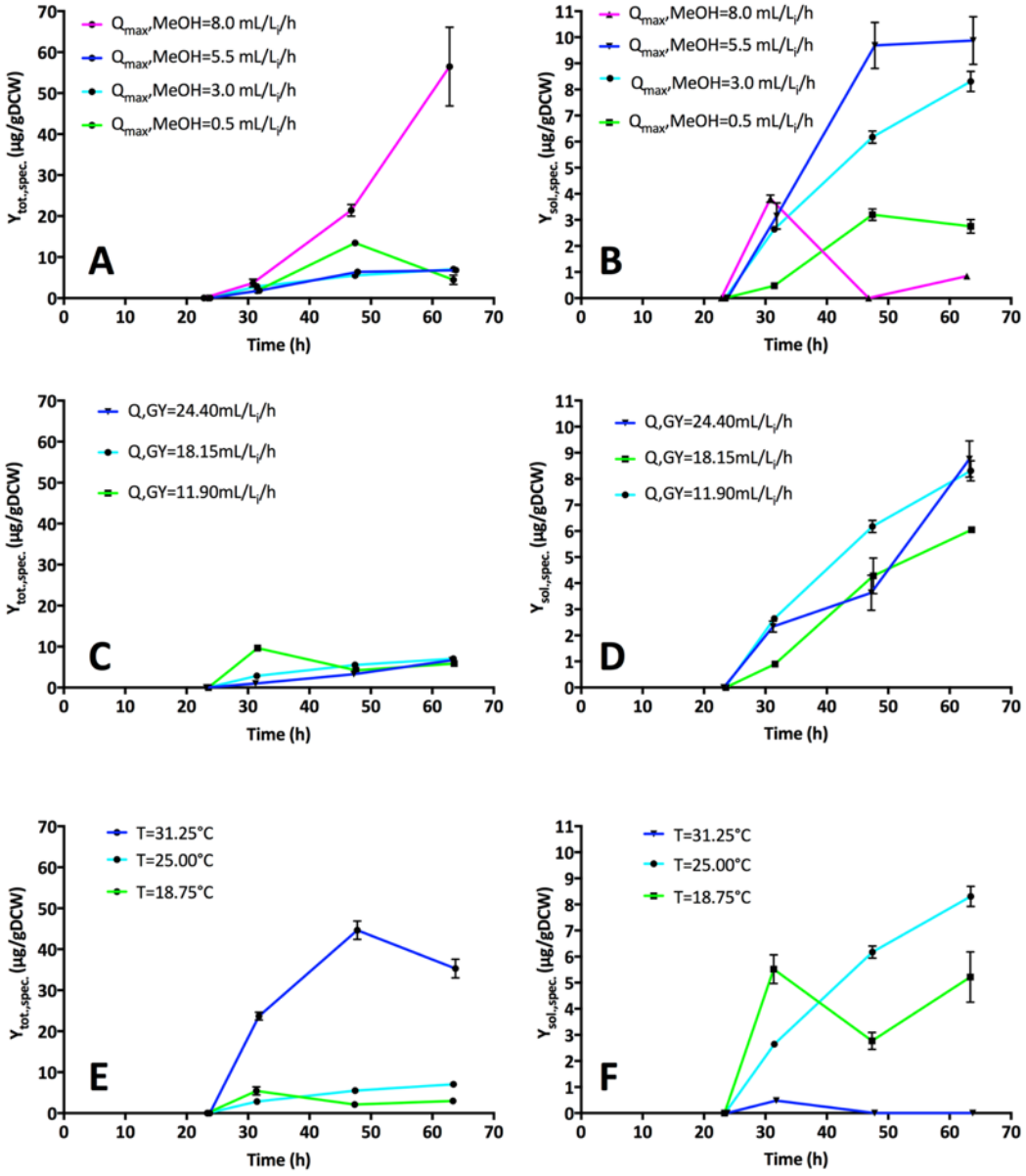

Figure 6.11 (A and B) Expression at various methanol flow rates of total and soluble tHBc respectively. (C and D) Expression at various glycerol flow rates of total and soluble tHBc respectively. (E and F) Expression at various induction temperatures of total and soluble tHBc respectively

6.5.4 Specific VLP-core protein material analysis

Material from the clarified supernatant was passed through a SEC column in order to identify material that is of the correct size range for the VLP. As shown in Figure 6.12, there are a number of differences between specific expression of VLP and soluble and/or total specific tHBc expression. Firstly, for VLP-fitting tHBc there is no apparent effect for the glycerol feed rate. Secondly, the induction temperature seems to have a significantly higher effect of specific VLP expression and also demonstrates an optimum value. Note that similar trends have been observed for volumetric expression of VLP protein (Figure 10.14).

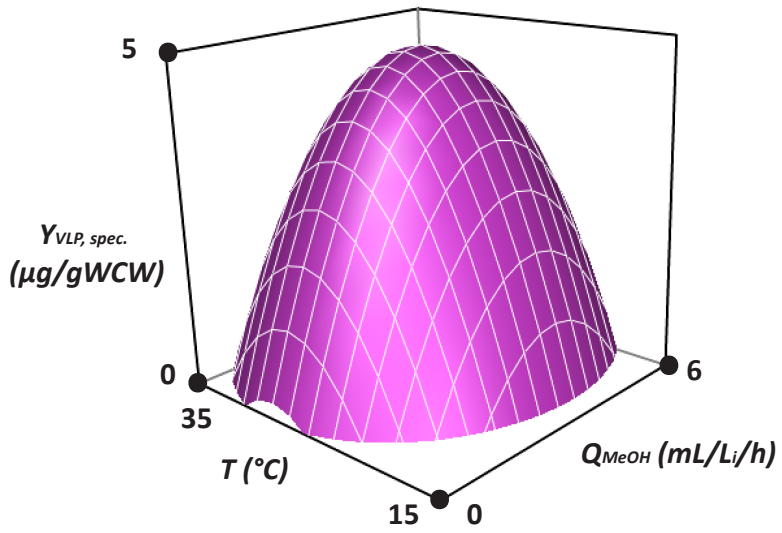

Figure 6.12 Response Surface Model for specific VLP core protein expression, $Y_{VLP, sol}$. Effect of Q_{MeOH} and T. Effect of Q_{GY} was found to be negligible

To understand VLP assembly a correlation between soluble tHBc and the proportion of soluble tHBc assembled into VLP-size is shown in Figure 6.13.

When looking at this figure, it appears that the higher the expression of soluble tHBc is, the lower the proportion of assembled VLPs to soluble tHBc is. Just like

several findings in chapter 3, supporting this data is the finding that VLP self-assembly requires regulation, as relatively high and uncontrolled expression of free units of tHBc can result in kinetic traps and thus incomplete formation of capsids (Zlotnick and Mukhopadhyay, 2011). Note that some degree of caution must be taken when interpreting the data in Figure 6.13. This is because a VLP-to-soluble tHBc ratio above 1 is unrealistic. The reason the existence of ratios higher than 1, is most likely due to a non-specific binding effect of HCPs during immunoblotting detection. Hence, there would be less non-specific binding in relatively pure samples (i.e. SEC-derived sampled) resulting in higher levels of detection in an immunoassay.

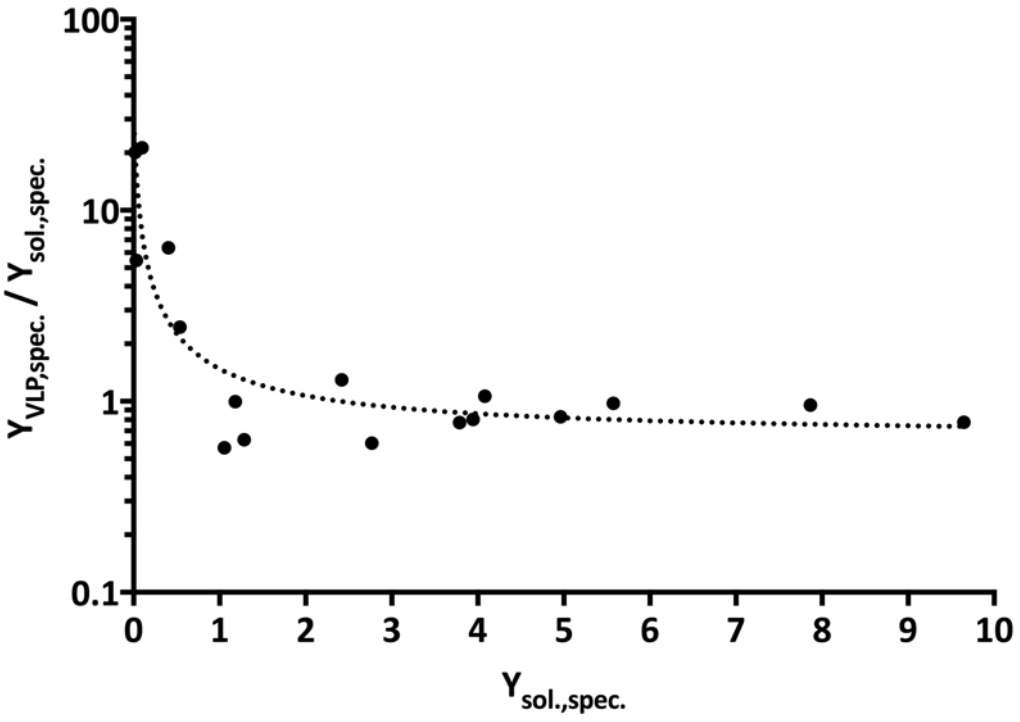

Figure 6.13 Correlation between specific soluble expression of tHBc (x-axis) and the relative proportion of soluble tHBc assembled into VLP-sized particles (y-axis). Each point in represents harvested material of the 17 fermentation runs discussed in this chapter.

Overall, it seems that there is a very small window within which soluble Tandem Core protein successfully assembles into VLPs per unit of biomass. One way to address this limitation would be minimise specific expression of tHBc while maximising the concentration of biomass. Lowering the induction temperature and methanol flow rate has been shown to achieve this effect. However, such a strategy would drastically increase downstream processing load as relatively more material would have to be removed to purify VLPs.

6.6 Response-specific optima

As seen in previous sections, different optimum settings are appropriate depending on the response studied. These optima are summarised in Figure 6.14.

Looking at Figure 6.14A, when trying to maximise biomass and metabolic activity, indicated by [X] and CER respectively, higher methanol flow rates are preferable (Q_{MeOH}=5.5mL/L_i/h). However, when aiming to maximise specific expression of tHBc, regardless of its state, a slightly lower flow rate is required (Q_{MeOH}~3.5mL/L_i/h). This setting is slightly higher when aiming to maximise volumetric expression since, as mentioned previously, this output is also a function of biomass.

Figure 6.14B shows the different glycerol flow rate setting required to achieve different optima. Unsurprisingly, to maximise biomass at the end of the induction phase the biomass at the start of the induction phase has to be maximised, as indicated by the highest possible feed setting within the range investigated. The opposite is seen when aiming to maximise the carbon evolution rate and minimising

the respiratory quotient. Glycerol flow rate optima were found to be at the design midpoints for all expression responses, except for VLP-fitting tHBc. In the latter case, the glycerol flow rate was not found to have a significant effect.

Figure 6.14C shows that to maximise biomass and metabolic activity, lower temperatures are preferable. For the lowest possible respiratory quotient, an optimum operating temperature of 22°C was established. Interestingly, the maximum investigated temperature was found to be the optimum for both total and soluble specific expression of tHBc. However, for maximum specific expression of VLPs, a significantly lower temperature was required. This again suggests, that for the production of VLPs, more control is required than simply upregulating expression of heterologous protein.

There are therefore trade-offs to be made when determining the fermentation protocol for the production of Tandem Core VLPs in *P. pastoris*. Lower temperatures seem more favourable for increasing the concentration of biomass. This, however, lowers specific expression but, in turn, seems to promote the specific expression of soluble Tandem Core material. On the other hand, high specific expression of Tandem Core protein, regardless of its state, is achieved at higher induction temperatures. A downside of this approach is a decrease in biomass and product solubility. For the purpose of this work, the aim was to maximise the volumetric production of VLPs. Optimum settings for this were an induction temperature of around 25°C and a methanol feed rate above 4.5 mL/L_i/h.

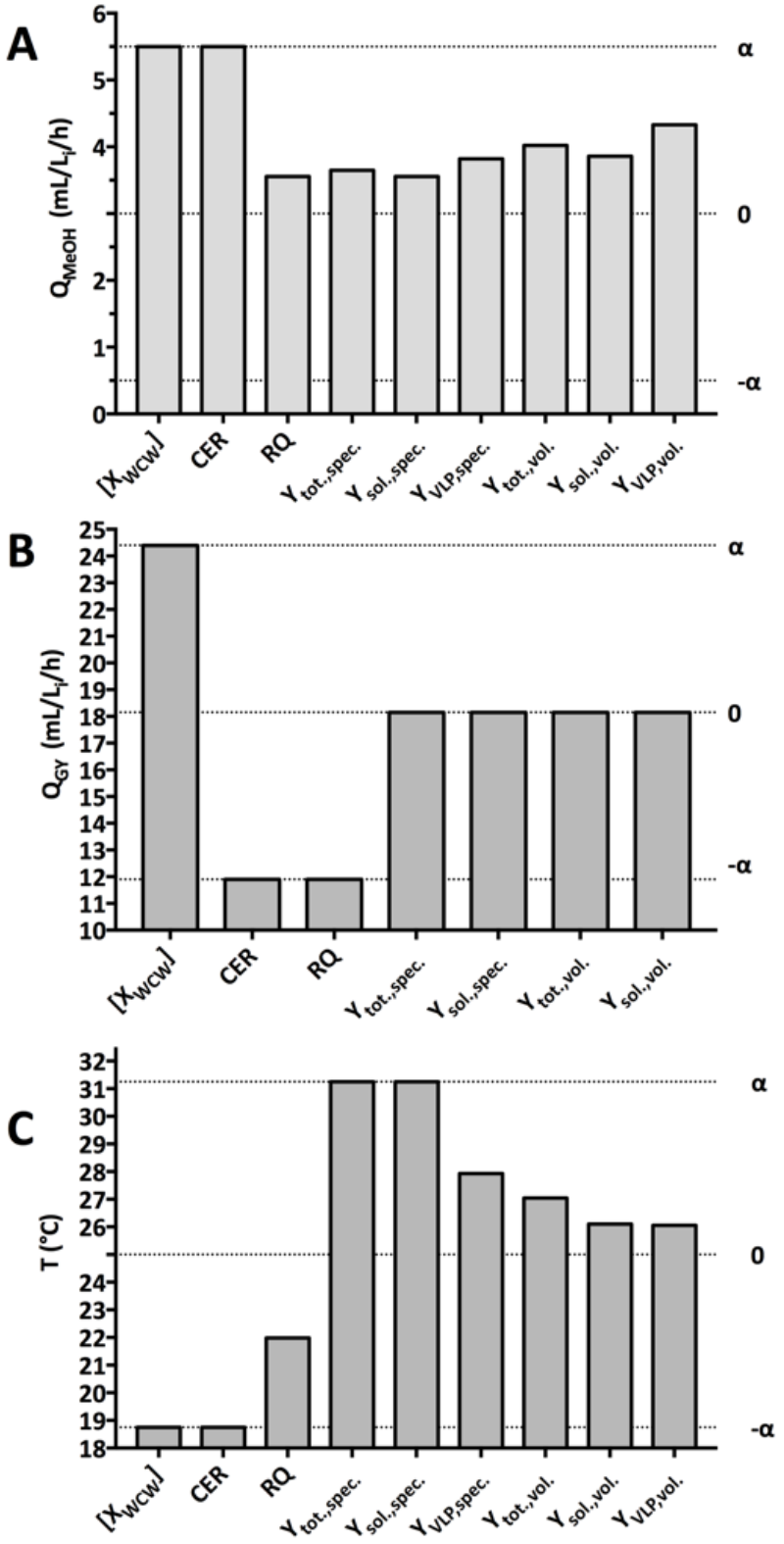

Figure 6.14 Summary of response optima for (A) methanol flow rate settings, (B) glycerol flow rate settings and (C) induction temperature settings. The symbols α and $-\alpha$ represent the centre face expansion coordinates of the design space, whereas 0 represents the midpoint coordinates.

6.7 Summary

Through time course analysis and response surface methodology the production of the k1,k1 construct in *P. pastoris* has been characterised. This characterisation yielded several key findings.

Metabolic responses, [X], CER and RQ, required significantly different optimum operating conditions than specific expression responses and volumetric responses. Assuming these models are not highly specific to the protein of interest being expressed, such data could translate to many processes employing *P. pastoris* to produce heterologous protein in general.

Methanol toxicity was studied and it was found that induction feed rates can be set significantly higher than recommended by industry standard protocols for these constructs.

Analysis of specific expression responses showed that models for expression of total and crude tHBc were very much alike. However, the model for specific VLP-assembled tHBc was very different. This indicated that maximising expression of tHBc was not synonymous to maximising the production of VLPs. It was, in fact, shown that a relatively narrow window for specific tHBc expression exists in which a significant part of Tandem Core units assembles into Virus-Like Particles. There are several suggested theories as to why tHBc expression and VLP assembly are not directly proportional. These are:

1. Higher expression rates could result in kinetic traps, resulting in unassembled intermediate material.
2. Expressed material is expressed disproportionately high to chaperone proteins, resulting in expression of misfolded and truncated Tandem Core units, which are either insoluble or incapable of assembling into a VLP.
3. Stress responses due to accumulation of heterologous protein as a result of internalisation.
4. Stress responses due to starvation or intoxication of culture.

Finally, volumetric production of VLPs was studied. This response was the most relevant response for applied vaccine production and was used a base for process scale-up. This will be covered in the next chapter.

6.8 Discussion

Aside from the metabolic responses, it should be stressed that, due to low r^2 values, some of the models displayed in this chapter cannot serve in a predictive fashion but, rather indicate several principle differences discussed previously, including the difference between tHBc expression and VLP assembly.

The noise of expression models could be significantly reduced by:

1. Using more sensitive assays.
2. Including sample process automation to rule out manual errors

3. Run all fermentation processes parallel instead of in campaign mode. Doing so would rule out any issues related to storage and age of pellets

In addition, future work should address the following points:

1. Harvest time should be included in all the response models.
2. Rates of tHBc expression and VLP assembly could provide more meaningful insight to the assembly process than concentrations at fixed time points.
3. Response surface methodology for residual methanol should be included to identify factor interactions affecting substrate consumption and induction stimuli.
4. RNA measurements would provide a better measure of heterologous protein expression than measuring tHBc concentrations in crude lysate.

However, in summary, this approach has increased our understanding of the design space for VLP production in *P. pastoris*, and permits for fine tuning of the fermentation depending on the desired optimum response.

Chapter 7 Scale-up and application of Tandem Core technology to universal influenza vaccine candidates

7.1 Introduction

Thus far, this work has characterised the production of Tandem Core VLPs using the k1, k1 VLP construct; and at small-scale, using the ambr®250 modular bioreactor system.

For this research to attain practical and industrial relevance, in particular with the aim of creating a universal influenza vaccine, these findings need to be translated to VLP production (1) with Tandem Core VLPs presenting influenza epitopes and (2) at larger scales.

7.2 Aims and objectives

This chapter aims to investigate how the previous findings translate to applied industrial scenarios. Thus, the following objectives were set:

- Investigate effects of varying inserts: This chapter will investigate the effects on fermentation processes and product yields as a result of changing epitopes on the Tandem Core VLP exterior.
- Investigate effects of scale-up: To attain industrial relevance it must be demonstrated that the established processes can be scaled-up.

- Investigate the effects of alternative feeding strategies on yield and scalability.

7.3 Variation of inserts with universal influenza epitopes

The k1, k1 VLP constructs utilises lysine residues in the major insertion regions (MIR) of the Tandem Core sub-unit. However, this work seeks to replace the lysine residues with small conserved regions of influenza that might be useful in a universal vaccine. Thus far, the effect of changing the insert on VLP production and assembly has yet to be characterised and therefore, this section aims to scope the effects of changing inserts on the Tandem Core scaffold on upstream process requirements. This was done by inducing various transformants of *P. pastoris* Mut^S expressing different Tandem Core constructs at identical fermentation conditions and subsequently, studying metabolic responses and product titers.

7.3.1.1 Pure methanol induction yields lower observable titres for more complex inserts

Several constructs were created to monitor the effect on upstream processing. These constructs are summarised below.

- **k1, k1**: As mentioned earlier, this represents the most simplified VLP construct, containing only single lysine inserts in the MIRs. This construct has also been investigated in the previous chapter.

- **LAH3, k1**: This construct consists of a lysine insert in one of the MIRs on the scaffold and a variant of Lower Alpha Helix (LAH) of the hemagglutinin stalk. Previous research has shown that this epitope provides the desired immunogenic response and can be placed onto a hepatitis B core VLP scaffold (Krammer and Palese, 2013; Chen *et al.*, 2015).
- **HA2.3, $(M2e)_3$**: This construct displays two different influenza-specific epitopes. The first epitope is HA2.3, which comprises the long-alpha helix structural domain, plus an extended amino acid sequence (Barbey-Martin *et al.*, 2002). $(M2e)_3$ represents a triplet sequence of an evolutionarily conserved region ectodomain of the M2 channel in influenza viruses. Previous studies have shown induced M2e-specific CD4 T immunity when M2e was presented on a VLP (Deng *et al.*, 2015). This section will investigate two *Pichia* transformants: one with a relative high copy number ('HC'); and one with a relatively low copy number ('LC')[17].
- **LAH3, $(M2e)_3$**: This construct displays the previously discussed epitopes $(M2e)_3$ and LAH3.

Using a pure methanol induction protocol (section 2.3.3.1), at 30°C, Q_{MeOH}=5.0 mL/L_i/h and Q_{GY}=13.15 mL/L_i/h, no significant differences in final biomass at 40 hours induction were observed between transformants (Figure 7.2). Minor differences between CER and RQ were observed Figure 7.1. These variances could relate to the differences in product stability, meaning less stable products could be

[17] These copy numbers were determined qualitatively and were based on differences in antibiotic resistance provided by the vectors used for transformation (data not shown).

recycled as a carbon source, or differences in the copy numbers of heterologous genes of these transformants, as demonstrated by two transformants containing the same HA2.3, $(M2e)_3$ construct with different copy numbers in Figure 7.1 (light blue vs dark blue). However, substantial differences between specific expression of VLPs were observed in Figure 7.2. It was found that transformants expressing single and relatively smaller inserts (k1, k1 and LAH, k1) had the highest specific yields. However, it should be noted that the chosen induction method was based on maximising total specific tHBc expression, rather than maximising total specific VLP production as the assembly kinetics for each novel VLP was unknown.

Therefore, it was postulated that assembly of Tandem Core protein with larger inserts would be more challenging at relatively high total expression rates than for Tandem Core protein with relatively small inserts.

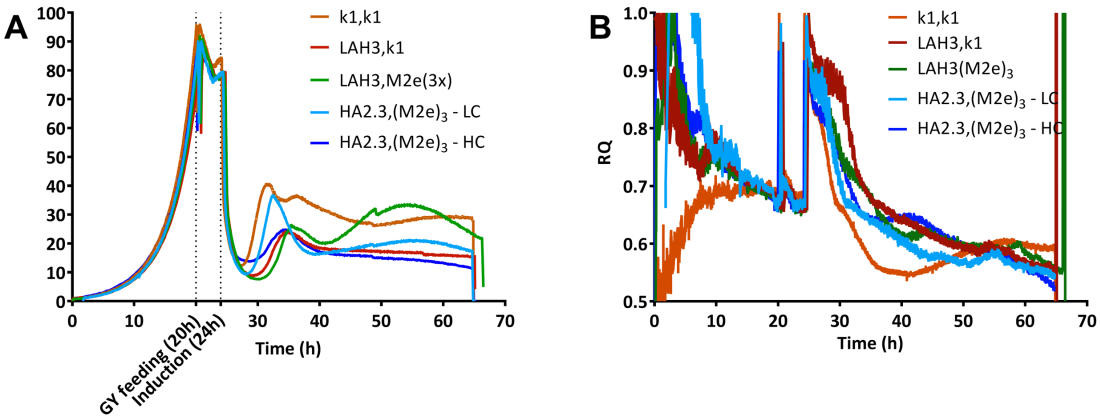

Figure 7.1 Representative metabolic anlaysis of different VLP constructs: (A) Carbon Dioxide Evolution Rates and (B) Respiratory Quotients.

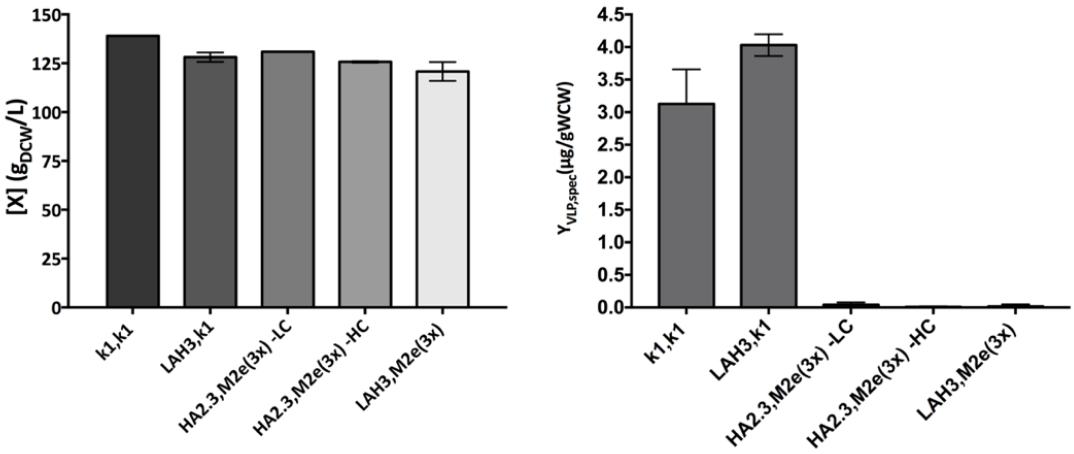

Figure 7.2 Quantitative analysis of harvested material of different VLP constructs: (Left) Biomass after 40 hours of pure methanol induction. (Right) Specific yields of different constructs.

This explains why the LAH3, k1 construct, containing an influenza epitope, was able to assemble into VLPs at comparable titres to the k1, k1 construct using the same induction conditions. Nonetheless, the aim of this research was to develop an influenza vaccine offering the broadest possible cross-protectiveness against influenza variants. Therefore, to maximise the potential of Tandem Core technology it was preferable to employ two inserts simultaneously, ideally using the HA2.3, $(M2e)_3$ construct. However, as shown by the data above this was not possible using a pure feeding strategy at these induction settings.

7.3.1.2 Mixed feeding induction provides better yields for more complex inserts

In an attempt to lower the specific expression of tHBc protein to facilitate the assembly of more complex Tandem Core units, a mixed feeding induction protocol was implemented in which glycerol was introduced simultaneously with methanol during the induction phase. The mixed feeding induction protocol is described in section 2.3.3.2. The aim of this strategy was (1) to create a carbon rich environment

such that cellular stress factors were not induced; and (2) to lower the specific expression of tHBc through partial repression of AOX genes and increased levels of biomass generation (Hartner and Glieder, 2006).

The HA2.3,$(M2e)_3$ (HC) construct was used for this study, as it provided the most complex version of insert combination available and would, therefore, according to the previously stated hypothesis, prove to be the most challenging version of Tandem Core protein to assemble into a VLP.

The data generated from this mixed feeding approach was compared to harvested material derived from pure methanol induced fermentation for the low-copy HA2.3,$(M2e)_3$ construct.

As shown in Figure 7.3A, a mixed feeding strategy resulted in higher metabolic activity as a result of higher biomass generation. As reported by several sources, this can be explained by the fact that the maximum specific growth rate is higher when glycerol is used as a carbon source than when methanol is used as a carbon source (Orman, Çelık and Ozdamar, 2009; Looser *et al.*, 2014). The utilisation of a different carbon source is confirmed through analysis of the RQ shown in Figure 7.3B, which shows consistently higher values throughout the induction phase for mixed feed induction than the pure methanol induction cultures. This, in turn, can be explained by the fact that the RQ for glycerol consumption (RQ_{GY}=0.86) is higher than for the consumption of methanol (RQ_{MeOH}=0.67).

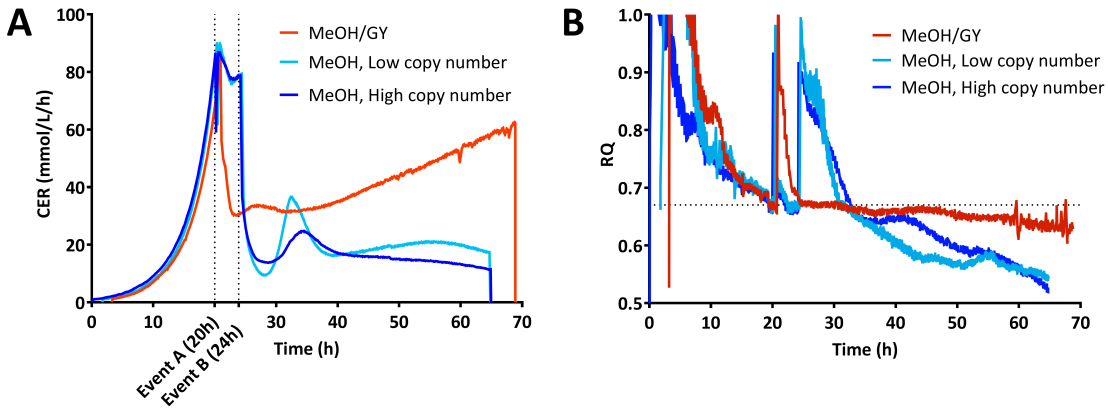

Figure 7.3 Metabolic analysis of HA2.3, $(M2e)_3$ constructs with different copy numbers at different induction modes. (A) Carbon Dioxide Evolution Rates. Event A shows the approximate time of the glycerol fed-batch phase for pure methanol feeding induction processes and the start of innduction for the mixed feeding induction process. Event B shows the time for the start of induction for the pure feeding processes. (B) Respiratory Quotients. The orange curves denote the responses of the mixed feeding induction process, whereas the blue curves denote the responses for the pure feeding induction processes at high copy numbers (dark blue) and low copy numbers (light blue).

Time course analysis of VLP-sized HA2.3, $(M2e)_3$ expression, shown in Figure 7.4, restated the earlier conclusion (Figure 7.2), that low titers are observed when using the pure methanol-based induction method described above. In fact, it was found that beyond 24 hours, a loss of specific productivity occurred. Furthermore, Figure 10.15 shows how the same pure methanol-based induction method was applied for the, simpler, k1,k1 VLP construct, yielding significantly different results. This suggests that VLP expression and assembly is highly dependent on the construct.

On the other hand, when comparing the titres from a mixed feed induction to pure methanol induction of the same VLP construct (HA2.3, $(M2e)_3$), specific titres are significantly higher throughout the entire induction phase. This implies the induction process has to be optimised for each individual VLP-construct.

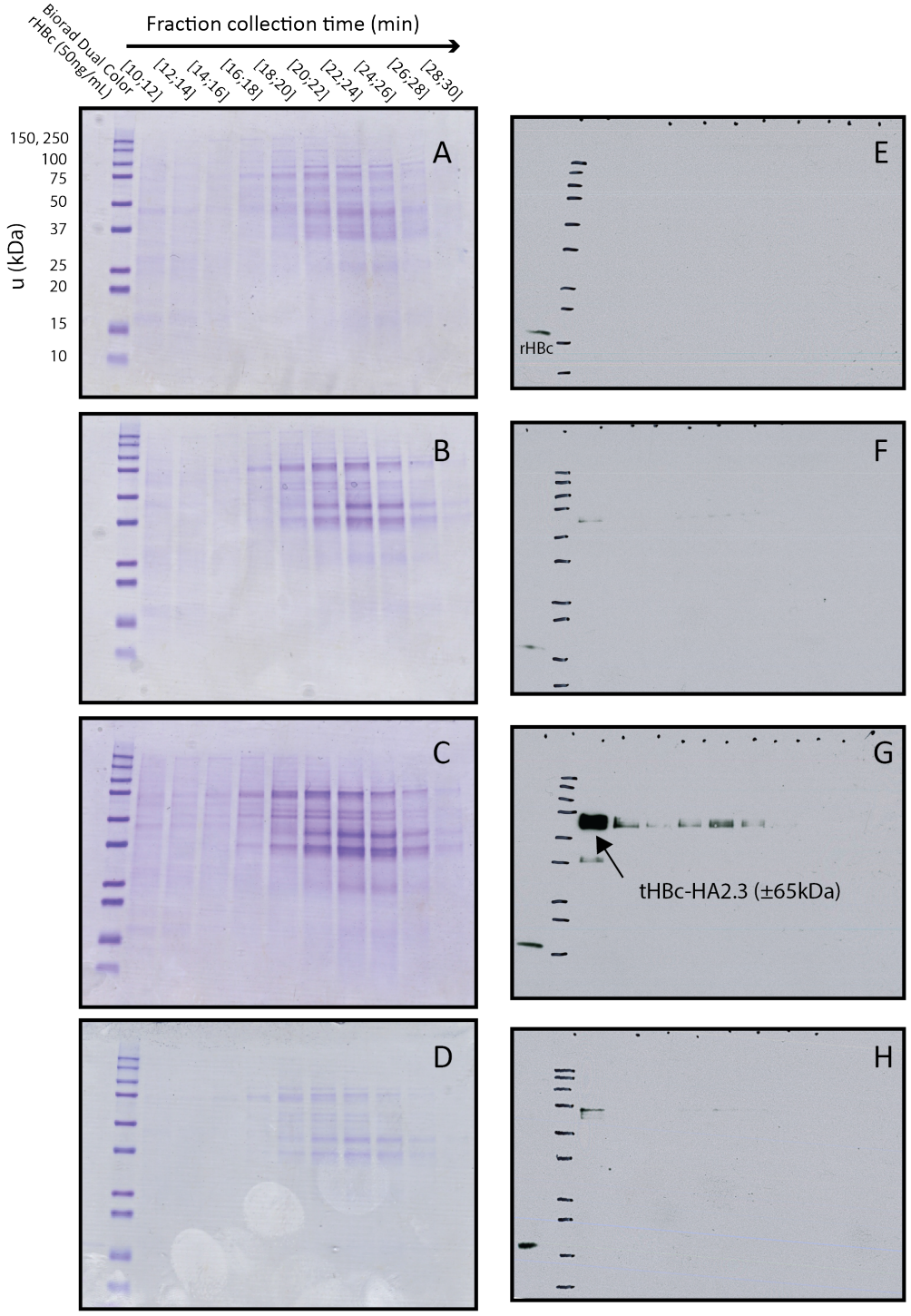

Figure 7.4 Size Exclusion Chromatography (2.4.11) fractions of bead-lysed (2.4.9) HA2.3(M2e)$_3$ construct ($\pm$ 65kDa) using pure methanol induction. Two-minute chromatography fractions were collected after 10 minutes for an additional 20 minutes. (A-D) Coomassie stains of reduced gels. (E-H) Western blots (1-minute exposure). Fermentation samples were collected at 0 hours (A & E), 8 hours (B & F), 24 hours (C & G) and 40 hours (D & H) of induction. SEC and Western blotting procedures were performed by iQur Ltd.

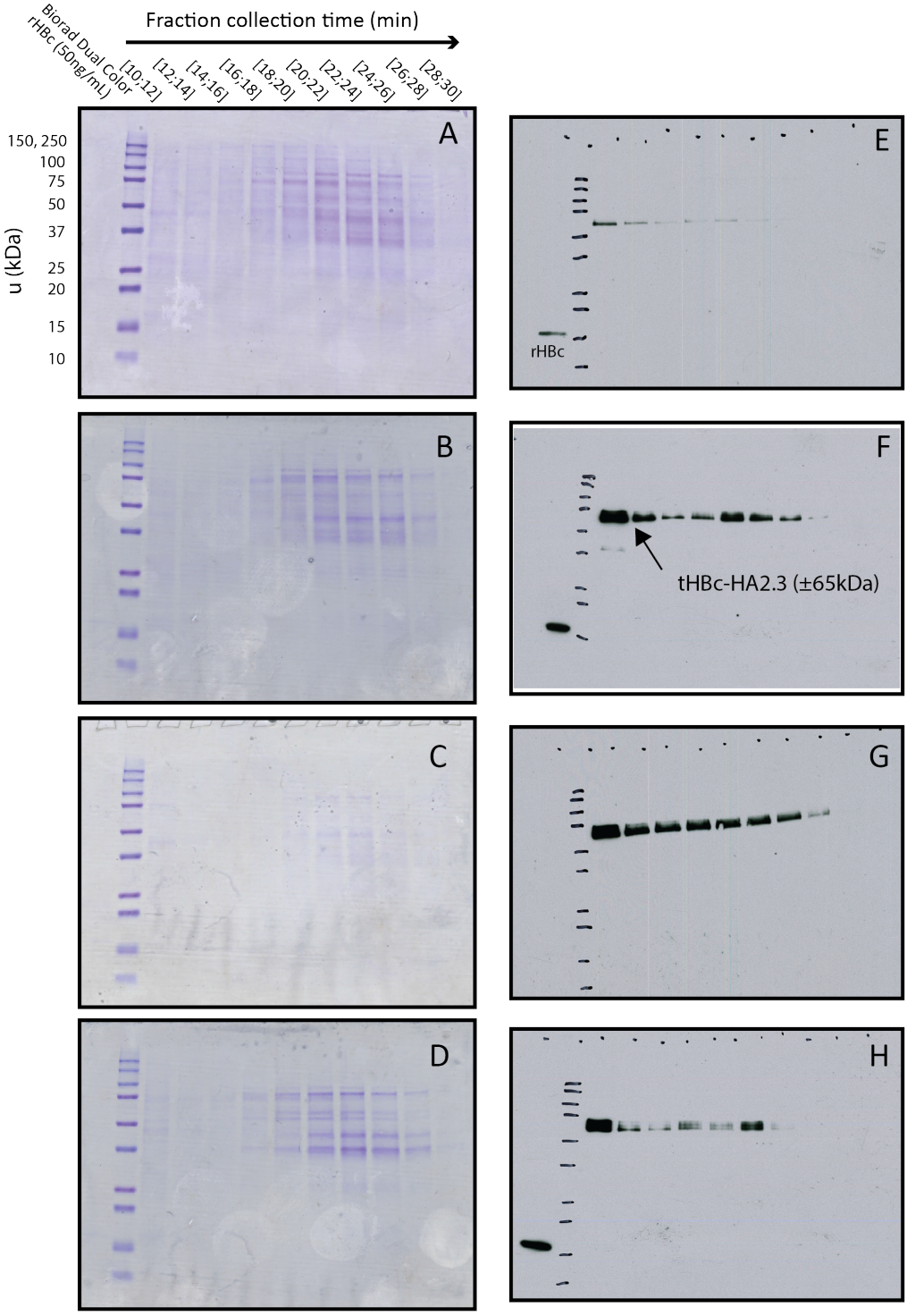

Figure 7.5 Size Exclusion Chromatography (2.4.11) fractions of bead-lysed (2.4.9) HA2.3(M2e)$_3$ construct ($\pm$65kDa) using glycerol mixed feed induction. Two-minute chromatography fractions were collected after 10 minutes for an additional 20 minutes. (A-D) Coomassie stains of reduced gels. (E-H) Western Blots (1-minute exposure). Fermentation samples were collected after 1.5 hours (A & E), 9 hours (B & F), 31 hours (C & G) and 48 hours (D & H) of induction. SEC and Western blotting procedures were performed by iQur Ltd.

7.4 Effects of scale-up

So far, this chapter has investigated two categories of induction modes at miniaturised scales of fermentation: pure methanol-based induction and glycerol/methanol mixed feed induction. This section will compare the scalability of each of these induction modes. The processes were scaled 120-fold: from miniaturised bioreactor scale (V_T=250mL) to pilot plant scale (V_T=30L).

Process scale-up was performed using constant power per volume input at constant relative gas flow rates. The power input is calculated as follows (Junker, 2004; Schmidt, 2005):

$$\boldsymbol{P = P_0 \rho N^3 D^5}$$ **Equation 7-1**

Where:

- P is the gassed power input (W)
- P_0 is the non-gassed power input (W)
- ρ is the fermentation broth density (kg/m^3)
- N is the impeller agitation speed (rpm)
- D is the impeller diameter (m^3)

Maintaining constant power per volume input is achieved through the following correlation:

$$\frac{\boldsymbol{P_1}}{\boldsymbol{V_1}} = \frac{\boldsymbol{P_2}}{\boldsymbol{V_2}}$$ **Equation 7-2**

Where the numbered subscripts are associated with different vessels.

For simplicity, geometric similarity was assumed. Hence, impeller agitation speeds at large-scale were determined though maintaining constant N^3D^2.

7.4.1 Challenging scale-up of pure methanol cultures

A high-yield process for VLP production of k1, k1-VLPs was established in the previous chapter and was chosen as the fermentation process to scale-up to investigate pure methanol induction scalability. The induction parameters for this scaled-down process were: T=25°C, Q_{GY}=13.15mL/L_i/h and Q_{MeOH}=5.0mL/L_i/h.

Figure 7.6A shows the near-identical off-gas carbon dioxide profiles during the batch phase and induction phase. However, during the glycerol fed-batch phase the observed carbon dioxide concentration in the exhaust gas was significantly higher at large-scale than at small-scale. An explanation for this is the fact that two different gas blending modes were used for the two scales. The addition of pure oxygen through gas blending typically comes into play during the glycerol fed-batch phase to maintain the increasing oxygen uptake rate of the culture. At small-scale, oxygen blending was performed through addition of oxygen through varying the total gas flow rate, whereas at large-scale, this mode of blending was not available. Instead, at large-scale, oxygen was blended while maintaining a constant gas flow rate. Because the gas flow rate was increased at small-scale, a lower gas concentration of carbon dioxide was observed. However, aside from different off-gas carbon dioxide profiles, the difference in gas blending modes also seems to have resulted in

different biomass concentration through the induction period. One possible reason for this is that the smaller scale bioreactors have reduced mass transfer capacity due to surface tension effects.

Figure 7.6C shows specific total and soluble tHBc yields over time. The corresponding fermentation samples were processed using the lyticase-augmented AFA method (sections 2.4.6 and 2.4.5) followed by dot blot analysis (sections 2.5.10 and 2.5.11). For the first 24 hours of induction, total specific expression seems to be highly similar between the two scales however, beyond this time, a drop in specific expression is seen at the larger scale. This could be explained by the fact that the larger scale fermentation attained higher biomass concentrations, hence more resources would have had to have been allocated to cell maintenance, rather than heterologous protein expression and this high biomass concentration would result in a lower specific expression, as one term is divided by the other.

Finally, volumetric productivity profiles were established (Figure 7.6D) and were found to follow similar trends to specific productivity profiles.

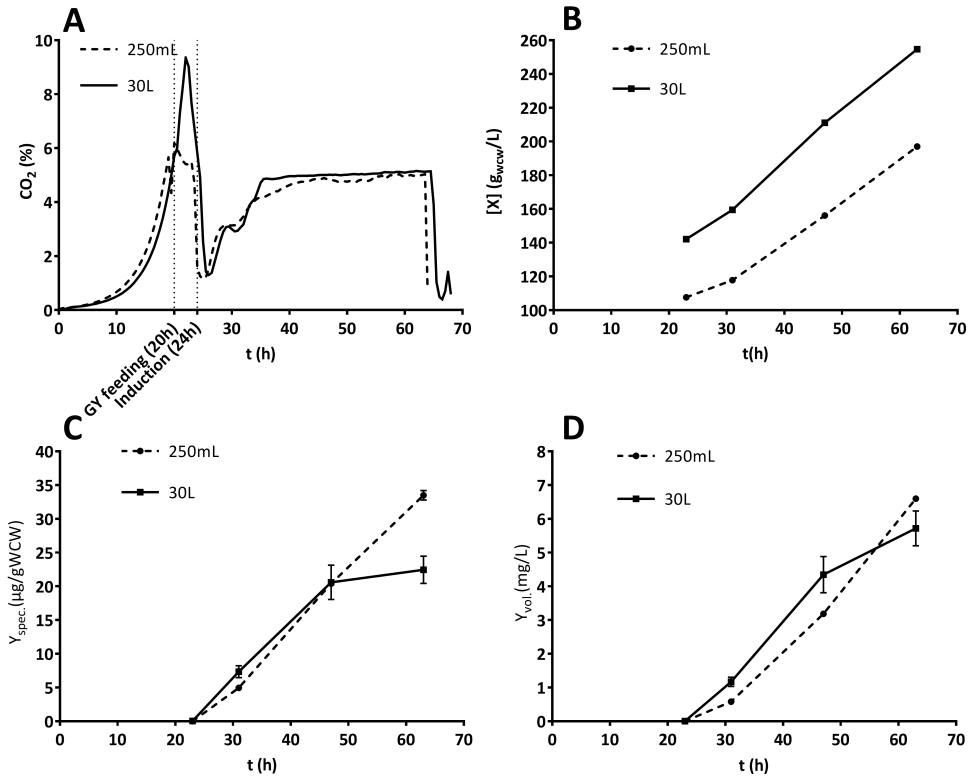

Figure 7.6 Fermentation data of process scale-up of *P. pastoris* expressing tHBc-k1,k1 using pure methanol-based heterologous protein induction. (A) Carbon dioxide concentrations measured in off-gas. (B) wet cell weight biomass concentrations. (C) Total specific expression of tHBc. (D) Total volumetric expression of tHBc.

7.4.2 Investigating scalability of mixed feeding induction

The previous section discussed the challenges observed when scaling up a pure methanol-based induction process. It was hypothesised that these differences arose during the glycerol fed-batch phase, where gas blending is most critical. On the other hand, a mixed feeding protocol omits this fed-batch phase and, therefore, did not require significant levels of pure oxygen gas blending.

It was hypothesised that scaling up the mixed feed induction process (section 7.3.1.2) would yield higher process similarity at large-scale than scaling up the pure methanol based induction process (section 7.4.1).

Figure 7.7 shows high similarities between both investigated scales in off-gas carbon dioxide profiles and biomass concentration profiles. Slightly lower values for both responses are observed at large-scale as the induction phase progresses. This could be explained by a slightly longer lag phase at the 30L scale, possibly caused by differences in inoculum culture concentration (data not shown).

Qualitative product analysis shown in Figure 7.8 shows near-identical expression profiles of total and soluble specific expressed tHBc in harvested material at 48 hours of induction. This similarity is emphasised when using log scaling Figure 10.16. This suggests that: (1) the studied mixed feeding induction process is highly scalable and, (2) it is more scalable than the investigated pure methanol induction process.

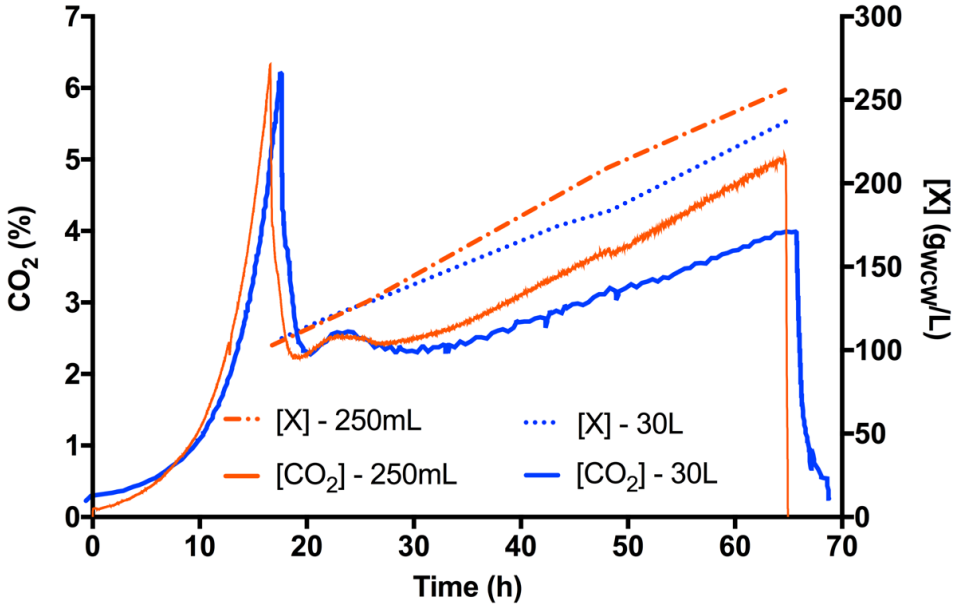

Figure 7.7 Fermentation data of process scale-up of *P. pastoris* expressing tHBc-HA2.3, $(M2e)_3$ using mixed feeding induction (section 2.3.5.2). Large-scale data, at 30L total working volume, is shown in blue. Small-scale data, at 250mL total working volume, is shown in orange. Carbon dioxide concentrations ([CO_2]) was measured in off-gas and biomass concentrations ([X]) were determined through wet cell weight analysis (section 2.5.2).

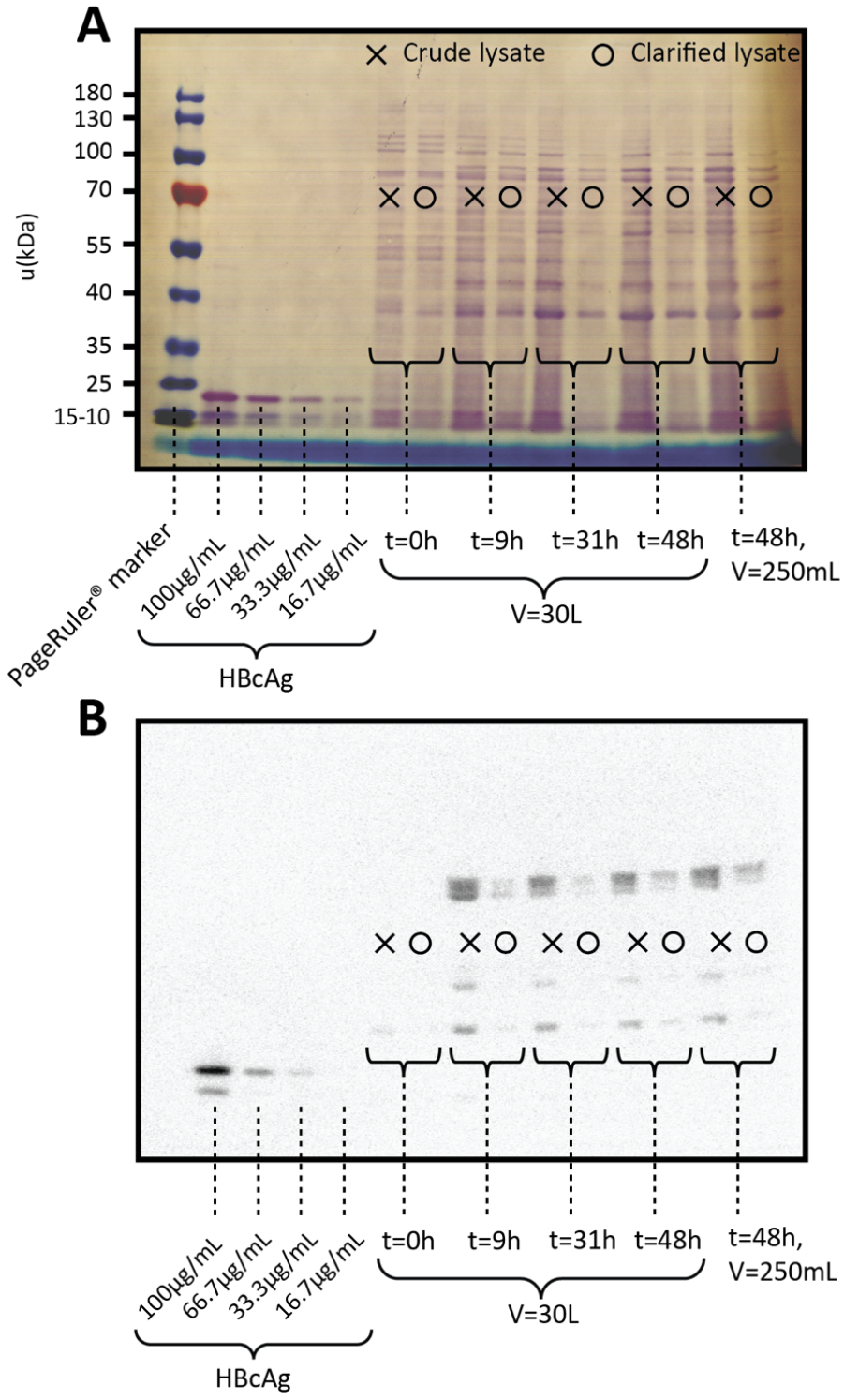

Figure 7.8 Qualitative protein data of large-scale and small-scale fermentations, both using the same mixed feeding induction protocol. For the large-scale fermentation (V_T=30L), samples were taken at 0, 9, 31 and 48 hours of induction. Samples for the small-scale fermentation (V_T=250mL) were taken at 48 hours of induction. Extractions were performed with homogenisation. (A) Membrane stain of reduced SDS-PAGE gel showing host cell protein expression profiles in crude and clarified lysates. (B) Western blot showing heterologous protein expression profiles of tHBc-HA2.3,$(M2e)_3$ ($\pm$65kDa) in crude and clarified lysates.

7.5 Impact of varying inserts on large-scale fermentation

The previous sections in this chapter have discussed the effects of variation of the epitopes inserted on the MIR of the VLP scaffold on product titres and other bioprocess-related responses, such as metabolic activity and respiratory quotient profiles; the suitability of two different modes of induction for the expression of different VLP constructs; and the process-scalability of these two different modes of induction. In comparison to pure methanol based-induction, it was found that mixed feeding induction was more suitable for VLP-constructs with more complex inserts, such as tHBc-HA2.3, $(M2e)_3$ and resulted in a more scalable process.

The final objective of this thesis was to use these findings in such a way that these would have a practical application. Therefore, this section aimed to combine the findings by investigating the effects of titres of different VLP constructs containing different combinations candidate influenza epitopes, produced at large-scale using the most scalable induction method, which, as discussed earlier was found to be mixed feeding. The candidate VLP constructs for a universal influenza vaccine were: (1) k1, LAH3; (2) k1, $(M2e)_3$ and (3) HA2.3, $(M2e)_3$.

Figure 7.9 shows that the differences between metabolic responses are relatively small when changing inserts on the MIR of VLPs when employing a mixed feed induction process.

However, expression profiles shown in Figure 7.10 were very different between the VLP constructs. This figure suggests that as the cumulative insert size increases, the

specific yield of VLP-fitting tHBc decreases. Figure 7.11 confirms this finding and shows the quantitative analysis of VLP yields. Figure 7.10 and Figure 7.11 also show that the investigated mixed feeding processes do not achieve as high yields as the optimised pure methanol-based induction process used to produce k1, k1.

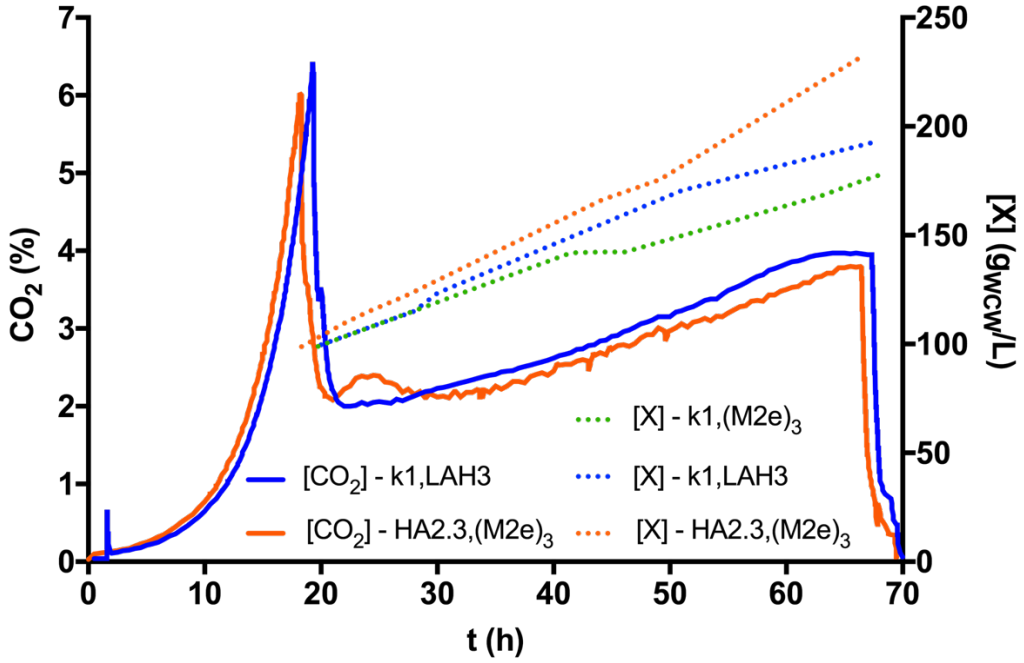

Figure 7.9 Carbon dioxide concentration profiles in fermentation exhaust gas and biomass concentration profiles for k1,LAH3, k1,(M2e)$_3$ and HA2.3,(M2e)$_3$ constructs during 30L mixed feed induction. Unfortunately, no carbon dioxide concentration profiles were obtained for the k1,(M2e)$_3$ construct. A component of these results has previously been published (Kazaks *et al.*, 2017)**.**

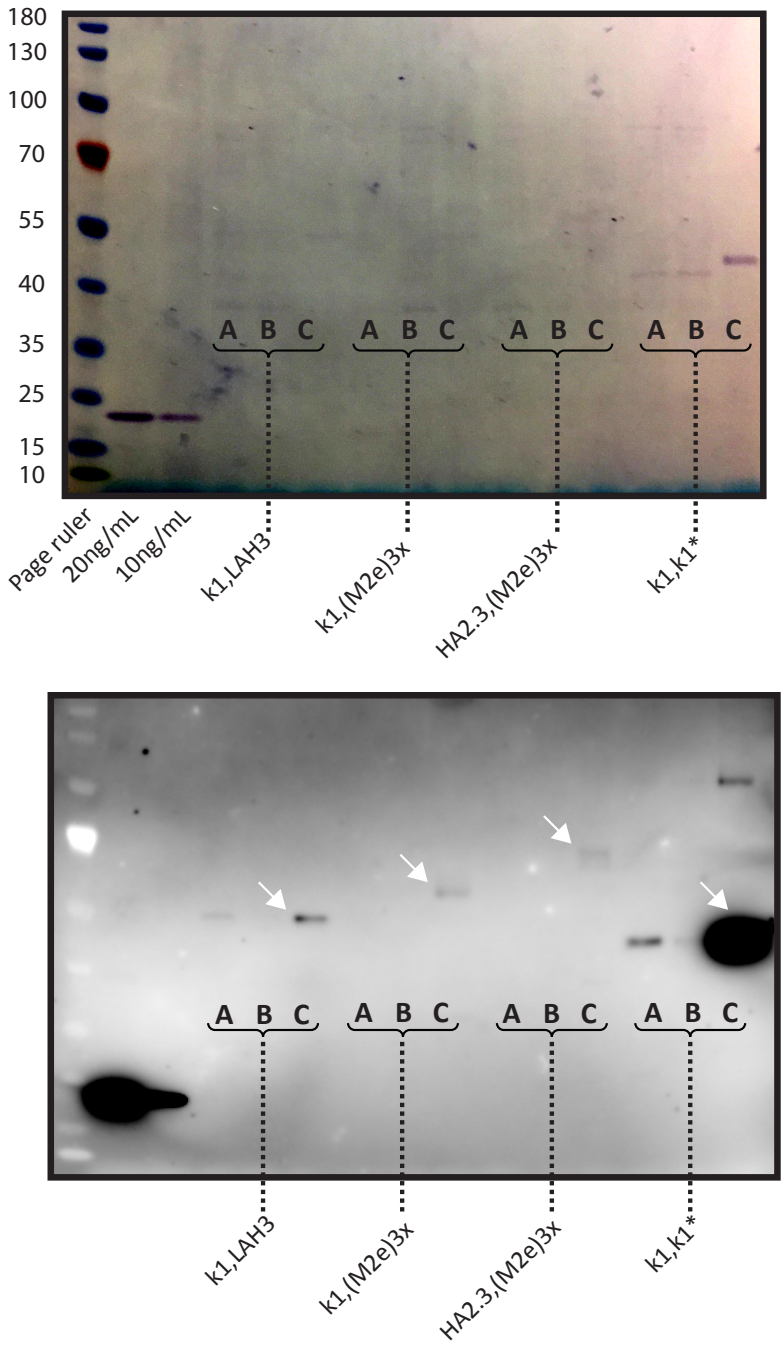

Figure 7.10 Qualitative analysis of expression of different tHBc constructs showing VLP-sized Tandem Core material through western blot analysis as described in sections 2.5.8, 2.5.9 and 2.5.11. All samples were standardised to a concentration of 50µg/mL total protein. Fermentations for the k1,LAH3 (±48kDa); k1,$(M2e)_3$ (±58kDa) and HA2.3,$(M2e)_3$ (±65kDa) were performed at 30L total volume and using a mixed feeding induction protocol. The k1,k1 construct (±42.5kDa); represents the yields of Tandem Core material produced through the pure feeding induction process as described in section 7.4.1. (Top) Reverse stained membrane of SDS-PAGE gel showing host cell protein expression profiles in crude lysate (A), clarified lysates (B) and VLP-sized material fractions (C). (Bottom) Western blot (1-hour exposure) showing heterologous protein expression profiles in crude lysate (A), clarified lysates (B) and VLP-sized material fractions (C).

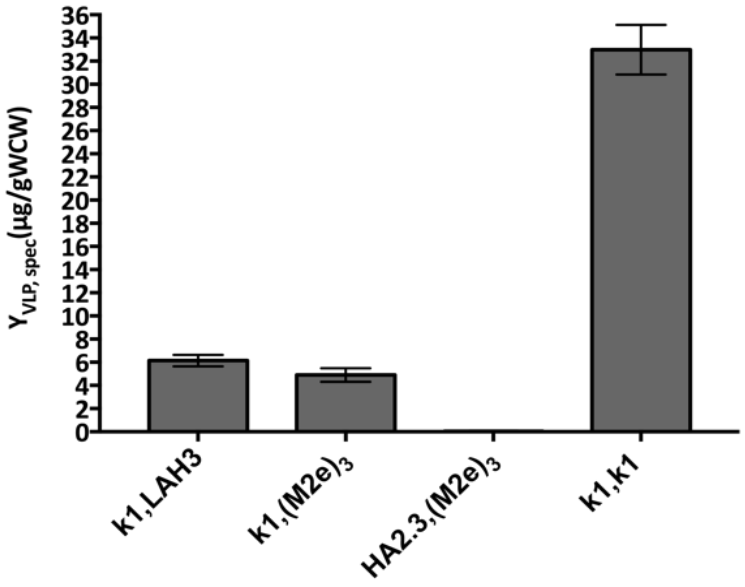

Figure 7.11 Quantitative analysis of expression of different tHBc constructs showing VLP-sized Tandem Core material through dot blot analysis as described in section 2.5.10. Fermentations for the k1,LAH3; k1,$(M2e)_3$ and HA2.3,$(M2e)_3$ were performed at 30L total volume and using a mixed feeding induction protocol. The k1, k1 construct represents the yield of VLP-sized Tandem Core material produced through the pure methanol induction process as described in section 7.4.1. Note that dot blot analysis was performed using linear fitting (r^2=0.96, n=18) rather than four parameter logistic fitting. This was done because 4PL analysis could not capture the higher-end concentration of the harvested and processed material from the k1,k1 fermentation. This could result in slight over-estimation of lower-end sample concentrations and slight underestimation of higher-end sample concentrations.

7.6 Summary

This chapter has demonstrated that there is no universal fermentation process for the production of the investigated VLP constructs: different fermentation processes must be investigated and established for different VLP constructs. It seems that the implementation of more complex inserts on the VLP exterior requires lower specific total tHBc expression. It was postulated that, aside from lowering induction temperature, this could be achieved by employing a glycerol mixed feed induction process. Indeed, for an influenza vaccine, providing broad cross-protection against influenza variants, larger and more complex inserts are required. It was

demonstrated that higher VLP titres containing these inserts were achieved using mixed feed induction when compared to the previously studied pure methanol-based induction processes.

Overall, the results initially demonstrated in the ambr® 250 system have been confirmed at pilot scale using the 30L fermenter and have netted similar results.

The scale-up of an optimised pure methanol-based induction process proved, to some degree, problematic. The issues that arose mainly resulted from the requirement of oxygen gas blending during the glycerol fed-batch phase. Because mixed feed induction omitted this phase in its process, the resulting process scale-up was much more successful.

Using this mixed feeding strategy, the expression of three candidate Tandem Core VLPs presenting various influenza epitopes, was investigated at pilot scale. It was again shown that the implementation of more complex inserts resulted in lower VLP titres. To fully understand why will require further study, but insert complexity may have a role to play.

Comparing the yields of these mixed feeding induction fermentations with the aforementioned (and poorly scalable) optimised pure methanol based induction process suggested that there was significant room for process improvement.

Chapter 8 Conclusions and future work

8.1 Conclusions

This chapter summarises the main findings of this thesis with respect to the specific aims identified in section 1.10.

Chapter 3 showed that, by lowering expression rates through changing the host cell phenotype and the introduction of smaller inserts of the Tandem Core scaffold, soluble and heat-stable VLPs were successfully produced in the appropriate size range. This chapter also identified a relatively large number of critical process parameters, therefore recognising the need for robust, high-throughput miniaturised fermentation platforms, more effective extraction processes; ideally combined with experimentally economic experimental design methodology.

The time and cost benefits of such miniaturised fermentation platforms could be gained only by employing complementary techniques facilitating high-throughput processing at small sample volumes. Chapter 4 therefore focused on the development of a high-performance, high-throughput, non-contact, automated, small-scale, scalable disruption tool for microbial bioprocess development. It was found that Adaptive Focused Acoustics (AFA) on enzymatically pre-treated samples provided the most suitable platform fulfilling these criteria. Development and optimisation of this method led to matching and even outperforming homogenisation performance, depending on the performance criterion investigated. The resulting overlap of performance ranges between High Pressure

Homogenisation and AFA-mediated cell disruption allowed for small-scale AFA performance mimicry of High Pressure Homogenisation. The utility of this work means that, for microbial cultures with sizeable cell walls, such as *P. pastoris*, an efficient scale-down tool is now available adding to the processing tools set.

It was found in chapter 5 that methanol-based induction of *P. pastoris* in microtitre plates proved to be challenging predominantly due to reduced mass transfer capacity and due to the tight regulation required of methanol induction. As an alternative, the use of pectin digest as an indirect induction agent was studied. It was found that the use of this novel media resulted in the expression of heterologous protein, however, it was determined that this effect was not due to the proposed metabolism of d-galacturonic methyl ester by pectin methyl esterase. Hence, it was found that microtiter plates provided an unsuitable platform for rapid upstream process development for methylotrophic induction. Nonetheless, several critical process parameters were confirmed: initial biomass concentration at induction, harvest time and methanol concentration in media.

The effects of variance of these critical process parameters were studied to a greater extent using miniaturised bioreactors in chapter 6. Metabolic responses, such as final biomass concentration, required significantly different optimum operating conditions than specific product expression responses and volumetric responses. This tightly controlled system allowed for a better understanding to be developed between tHBc expression and the formation of VLPs. It was found that maximising expression of heterologous protein was not synonymous to maximising

the production of VLPs. It was, in fact, shown that a relatively narrow window for specific tHBc expression exists in which a significant part of Tandem Core units assembles into VLPs.

Chapter 7 showed that different fermentation processes are appropriate for different constructs. It was demonstrated that higher titres for complex VLPs were achieved using mixed feeding induction rather than pure methanol-based induction. In addition, the scale-up of an optimised pure methanol-based induction process proved problematic. The issues that arose mainly resulted from the requisite of oxygen gas blending during the glycerol fed-batch phase. Because mixed feeding induction omitted this phase in its process, the resulting process scale-up was more successful. Using this mixed feeding strategy, the expression of three influenza vaccine candidates was investigated at pilot scale. It was again shown that the implementation of more complex inserts resulted in lower VLP titres.

8.2 Future work

Although chapter 3 showed how critical process parameters were identified associated with fermentation primary recovery, relatively little investigation was performed on the development of high-throughput analytical techniques that could distinguish between assembled and unassembled Tandem Core material. A main assumption in this thesis has been that Tandem Core-positive material within a particular size range represented assembled VLP. However, it must be stressed that this provides but an approximation as the method does not distinguish between

VLPs and Tandem Core-positive aggregates corresponding to the same size as a VLP. Future work should therefore focus on the development of alternative analytical techniques. Nonetheless, for preparative purposes, Size Exclusion Chromatography is a highly scalable purification technique and therefore has merit. It is therefore also recommended that the employed SEC method is improved to maximise recovery of VLPs.

The results in chapter 4 demonstrated that AFA-mediated cell disruption can be employed in high-throughput miniaturised fashion. This technology could therefore be used as a tool for a wide variety of applications including lysis buffer development, strain selection, fermentation process development and whole bioprocess integration. It is therefore recommended that future work focuses on broadening the range of applications for AFA. It should be noted that, in doing so, certain conditions could render lyticase less effective. Therefore, it is recommended that in-depth studies are dedicated to establishing the limitations of implementing lyticase pre-treatment in cell disruption processes of *P. pastoris*. Factors to be taken into consideration may include media pH, salt concentration, temperature, and/or various economic criteria; such as cost and throughput.

AFA has also demonstrated its effectiveness in cell disruption and product recovery in comparison to the industry standard of cell disruption, High Pressure Homogenisation, at small-scale. Cell disruption of *P. pastoris* using AFA could be implemented at industrial scale using continuous flow systems as pilot scale AFA

units now exists that could offer better recovery with less micronised debris. This could have significant implications for the development of continuous bioprocesses.

Chapter 5 focused on the development of microscale fermentation platforms. Although the experiments show that growth and induction can be achieved using microtitre plates, tighter process controls are required, particularly relating to the introduction of methanol into media. Future work should therefore focus on two areas: (1) the implantation of more sophisticated microscale fermentation platforms, such as the ambr® 15 micro bioreactor (Sartorius AG, Goettingen, DE) or the Micro-Matrix system (Applikon Biotechnology B.V., Delft, NL); and (2) further investigation of indirect AOX induction through the introduction of pectin oligomers.

Chapter 6 showed that tHBc expression and VLP assembly are not directly proportional. To prove and assess the effect of disproportionately high rates of tHBc expression in comparison to the required chaperone proteins, an adaptation to the methodology described by Damasceno *et al.* (2007) could be employed. This would involve the co-overexpression of chaperones such as immunoglobin binding protein (BiP) and Protein Disulphide Isomerase (PDI). Higher levels of co-expression could prove to be favourable for VLP assembly. In addition, cellular stress responses to either high levels of tHBc internalisation or high methanol concentrations could be correlated to stress-associated protein expression as described by others (Martin Dragosits *et al.*, 2009).

Although the number of factors in the experimental design was limited to three, response surface models could be generated where time, pH and DOT levels are included. Time was not included in the response models for VLP production as this would require large sampling volumes throughout the fermentations[18], and would significantly reduce bioreactor working volumes. However, if better analytical tools are developed to detect VLPs, these could be combined with the developed scale-down AFA-based cell disruption platform to eliminate the need for large samples volumes and to increase experimental throughput. Acidity was also not considered a priority factor as the trace salts the Basal Salts Media tended to precipitate above a pH of 5.5, significantly limiting the upper range of the experimental design space. However, there could be benefits in studying the effect of increased acidity as pH values below 3 could inhibit protease activity triggered by stress responses, thereby minimising the formation of truncated heterologous protein (Invitrogen Corporation, 2002). DOT was also not considered a priority factor, as an increase in DOT would mainly result in an increase in excess oxygen concentrations. Also, assuming a scenario in which high DOT levels are most favourable, maintaining relatively high DOT levels would become problematic when scaling up a fermentation process. Nonetheless, the effect of variance of DOT levels on heterologous protein expression by *P. pastoris* has reported to be significant and would be worthwhile to investigate (Invitrogen Corporation, 2002; Lee *et al.*, 2003).

Chapter 7 showed the benefits of implementing mixed feeding induction. A secondary consequence of adapting this approach would be the exclusion of the

[18] Due to the minimum sample volume required for preparative SEC

glycerol fed-batch phase. This would reduce the number of factors to be studied, making it more feasible to study the additional factors of interest discussed above. It should also be stressed that the scale-up of processes discussed in chapter 7 assumed geometric similarity of miniature and pilot scale bioreactors for the purpose of simplicity. In reality, there are dissimilarities between these vessel geometries. Therefore, to improve process scalability, accounting for the geometric differences between the vessels could account for better scale-up. Additionally, alternative scaling strategies could also be investigated, for instance, maintaining constant oxygen transfer capacity has proven to be an effective basis for scale-up when vessel geometry is dissimilar (Islam *et al.*, 2008).

This work has shown that upstream process development is highly dependent on the VLP construct involved. Future work could therefore explore the fundamentals of VLP assembly and its change with insert variation. However, the requirement of reassessing the suitability of a bioprocess after candidate epitope selection significantly impedes vaccine development. To move towards a more flexible vaccine development platform, a universal vaccine production process, such variance needs to be eliminated or reduced. This could be achieved by implementing a different VLP-production strategy altogether, in which a standardised VLP scaffold is produced in a single fermentation process followed by epitope conjugation in a separate unit operation post-purification. The investigated k1, k1 construct would in fact be suitable for lysine residue-based conjugation. Future work in this scenario would mainly involve optimising such a conjugation

step, which is arguably less complex and less time consuming than the optimisation of fermentation processes.

Finally, to improve volumetric yields it could be beneficial to develop transformants that secrete Tandem Core protein and prevent VLP assembly intracellularly so that VLP assembly can be controlled outside the cell in a separate unit operation. VLP assembly of hepatitis B core protein outside cells has indeed been demonstrated in previous works (Ludgate *et al.*, 2016). Such an approach would be beneficial because it would (1) reduce intracellular proteolytic degradation of tHBc resulting from a variety of possible cellular stress responses, (2) allow for highly controlled assembly of VLPs and (3) eliminate the need of a cell disruption method. The latter unit operation has been proven in this work to have the potential to significantly impede VLP recovery due to the generation of micronised debris. This secretory assembly could possibly be achieved by developing a construct that contains a fused SUMO-tag attached to the existing tHBc sequence. This SUMO tag would prevent VLP assembly until it is cleaved by specific proteases as described by others (Panavas, Sanders and Butt, 2009).

Chapter 9 References

Abate, T. (2013) *Stanford researchers take a step toward developing a 'universal' flu vaccine*. Available at: http://news.stanford.edu/news/2013/december/universal-flu-vaccine-121813.html (Accessed: 1 January 2017).

Alberts, B., Bray, D., Hopkin, K., Johnson, A., Lewis, J., Raff, M., Roberts, K. and Walter, P. (2004) 'From DNA to Protein: How Cells Read the Genome', in *Essential Cell Biology*. New York: Garland Science, Taylor & Francis Group, pp. 229–265.

Anand, H., Balasundaram, B., Pandit, A. B. and Harrison, S. T. L. (2007) 'The effect of chemical pretreatment combined with mechanical disruption on the extent of disruption and release of intracellular protein from *E. coli*', *Biochemical Engineering Journal*, 35(2), pp. 166–173. doi: 10.1016/j.bej.2007.01.011.

Anasontzis, G. E., Penã, M. S., Spadiut, O., Brumer, H. and Olsson, L. (2014) 'Effects of temperature and glycerol and methanol-feeding profiles on the production of recombinant galactose oxidase in *Pichia pastoris*', *Biotechnology Progress*, 30(3). doi: 10.1002/btpr.1878.

Barbey-Martin, C., Gigant, B., Bizebard, T., Calder, L. J., Wharton, S. a, Skehel, J. J. and Knossow, M. (2002) 'An antibody that prevents the hemagglutinin low pH fusogenic transition.', *Virology*, 294(1), pp. 70–4. doi: 10.1006/viro.2001.1320.

BioGrammatics, I. (2013) A short history of *Pichia pastoris*. Available at: https://www.biogrammatics.com/categories/pichia-pastoris-history (Accessed: 1 January 2017).

Bláha, B. A. F., Morris, S. A., Ogonah, O. W., Maucourant, S., Crescente, V., Rosenberg, W. and Mukhopadhyay, T. K. (2017) 'Development of a high-throughput microscale cell disruption platform for *Pichia pastoris* in rapid bioprocess design', *Biotechnology Progress*. doi: 10.1002/btpr.2555.

Blaha, B., Hughes, C., Khalid, V., Migani, D., Perales, L. R. and Umar, S. (2012) *Influenza Vaccine: A Design Brief for Future Manufacture of Seasonal and Pandemic Influenza Vaccine*. University College London

Brooks, G. F., Carroll, K. C., Butel, J. S., Morse, S. A. and Mietzner, T. A. (2013) 'Chapter 39. Orthomyxoviruses (Influenza Viruses)', in *Jawetz, Melnick, & Adelberg's Medical Microbiology, 26e*. New York City: McGraw-Hill, pp. 536–549.

Burden, D. (no date) 'Zymolyase ™ vs Lyticase and Glusulase'. Available at: http://www.amsbio.com/brochures/Zymolyase_Comparison.pdf.

Cann, A. J. (2012) 'Classification of Subcellular Infectious Agents', in *Principles of Molecular Virology*. Oxford: Elsevier, pp. 285–289.

Carnicer, M., Baumann, K., Toplitz, I., Sanchez-Ferrando, F., Mattanovich, D., Ferrer, P. and Albiol, J. (2009) 'Macromolecular and elemental composition analysis and

extracellular metabolite balances of Pichia pastoris growing at different oxygen levels', *Microbial Cell Factories*, 8(1), p. 65. doi: 10.1186/1475-2859-8-65.

CDC (2011) *CDC.gov*. Available at: http://www.cdc.gov/vaccines/pubs/pinkbook/downloads/prinvac.pdf (Accessed: 1 January 2014).

CDC (2018) *Influenza vaccines — United States, 2017–18 influenza season*. Available at: https://www.cdc.gov/flu/protect/vaccine/vaccines.htm.

Celik, E., Calik, P. and Oliver, S. G. (2009) 'Fed-batch methanol feeding strategy for recombinant protein production by *Pichia pastoris* in the presence of co-substrate sorbitol', *Yeast*. John Wiley & Sons, Ltd., 26(9), pp. 473–484. doi: 10.1002/yea.1679.

Cereghino, J. L. and Cregg, J. M. (2000) 'Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*', *FEMS Microbiol Rev*, 24. doi: 10.1111/j.1574-6976.2000.tb00532.x.

Chakrabarti, A., Singh, K., Narang, A., Singhi, S., Batra, R., Rao, K. L. N., Ray, P., Gopalan, S., Das, S., Gupta, V., Gupta, A. K. and Bose, S. M. (2001) 'Outbreak of', *Society*, 39(5), pp. 1702–1706. doi: 10.1128/JCM.39.5.1702.

Chanturiya, A. N., Basanez, G., Schubert, U., Henklein, P., Yewdell, J. W. and Zimmerberg, J. (2004) 'PB1-F2, an Influenza A Virus-Encoded Proapoptotic

Mitochondrial Protein, Creates Variably Sized Pores in Planar Lipid Membranes', *Journal of Virology*, 78(12), pp. 6304–6312. doi: 10.1128/JVI.78.12.6304-6312.2004.

Chen, A. K., Chen, X., Choo, A. B. H., Reuveny, S. and Oh, S. K. W. (2010) 'Expansion of Human Embryonic Stem Cells on Cellulose Microcarriers', in *Current Protocols in Stem Cell Biology*. Hoboken, NJ, USA: John Wiley & Sons, Inc. doi: 10.1002/9780470151808.sc01c11s14.

Chen, D., Nims, R., Dusing, S., Miller, P., Luo, W., Quertinmont, M., Parekh, B., Poorbaugh, J., Boose, J. A. and Atkinson, E. M. (2008) 'Root cause investigation of a viral contamination incident occurred during master cell bank (MCB) testing and characterization – A case study', *Biologicals*, 36(6), pp. 393–402. doi: 10.1016/j.biologicals.2008.07.005.

Chen, S., Zheng, D., Li, C., Zhang, W., Xu, W., Liu, X., Fang, F. and Chen, Z. (2015) 'Protection against multiple subtypes of influenza viruses by virus-like particle vaccines based on a hemagglutinin conserved epitope', *BioMed Research International*. Hindawi Publishing Corporation, 2015. doi: 10.1155/2015/901817.

Chen, W., Calvo, P. A., Malide, D., Gibbs, J., Schubert, U., Bacik, I., Basta, S., O'Neill, R., Schickli, J., Palese, P., Henklein, P., Bennink, J. R. and Yewdell, J. W. (2001) 'A novel influenza A virus mitochondrial protein that induces cell death', *Nature Medicine*, 7(12), pp. 1306–1312. doi: 10.1038/nm1201-1306.

Chiruvolu, V., Cregg, J. M. and Meagher, M. M. (1997) 'Recombinant protein production in an alcohol oxidase-defective strain of *Pichia pastoris* in fed batch fermentations', *Enzyme and Microbial Technology*, 21(4), pp. 277–283. doi: 10.1016/S0141-0229(97)00042-2.

Chowell, G., Nishiura, H. and Bettencourt, L. M. A. (2007) 'Comparative estimation of the reproduction number for pandemic influenza from daily case notification data', *Journal of The Royal Society Interface*, 4(12), pp. 155–166. doi: 10.1098/rsif.2006.0161.

College of Physicians of Philadelphia (2014) *History of Vaccines*. Available at: http://www.historyofvaccines.org/content/types-vaccines (Accessed: 1 January 2017).

Couderc, R. and Baratti, J. (1980) 'Oxidation of Methanol by the Yeast, *Pichia pastoris*. Purification and Properties of the Alcohol Oxidase', *Agric. Biol. Chem.*, 44(10), pp. 2279–2289. doi: 10.1271/bbb1961.44.2279.

Covaris (2011) 'E-series User manual'. Available at: http://covarisinc.com/wp-content/uploads/pn_010041.pdf.

Covaris (2017) *Covaris Product catalogue: microTUBE*. Available at: http://covarisinc.com/products/afa-tubes-and-vials/microtube/ (Accessed: 1 January 2017).

Cox, M. M. J. and Hashimoto, Y. (2011) 'A fast track influenza virus vaccine produced in insect cells', *Journal of Invertebrate Pathology*, 107, pp. S31–S41. doi: 10.1016/j.jip.2011.05.003.

Cregg, J. M. (2007) 'Distinctions Between *Pichia pastoris* and Other Expression Systems', *Pichia Protocols*, 389, pp. 1–10. doi: 10.1007/978-1-59745-456-8_1.

Cregg, J. M., Barringer, K. J., Hessler, A. Y. and Madden, K. R. (1985) '*Pichia pastoris* as a host system for transformations.', *Molecular and cellular biology*, 5(12), pp. 3376–85. Available at: http://www.ncbi.nlm.nih.gov/pubmed/3915774.

Crow, M., Deng, T., Addley, M. and Brownlee, G. G. (2004) 'Mutational Analysis of the Influenza Virus cRNA Promoter and Identification of Nucleotides Critical for Replication', *Journal of Virology*, 78(12), pp. 6263–6270. doi: 10.1128/JVI.78.12.6263-6270.2004.

Damasceno, L. M., Anderson, K. A., Ritter, G., Cregg, J. M., Old, L. J. and Batt, C. A. (2007) 'Cooverexpression of chaperones for enhanced secretion of a single-chain antibody fragment in *Pichia pastoris*', *Applied Microbiology and Biotechnology*, 74(2), pp. 381–389. doi: 10.1007/s00253-006-0652-7.

Deng, L., Cho, K. J., Fiers, W. and Saelens, X. (2015) *M2e-Based Universal Influenza A Vaccines*. doi: 10.3390/vaccines3010105.

Doran, M. D. (1996) *Bioprocess engineering principles*, *Minerals Engineering*. doi: 10.1016/S0892-6875(96)90075-8.

Doran, P. M. (2013) 'Cell Disruption', in *Bioprocess Engineering Principles*, p. 467.

Dragosits, M., Stadlmann, J., Albiol, J., Baumann, K., Maurer, M., Gasser, B., Sauer, M., Altmann, F., Ferrer, P. and Mattanovich, D. (2009) 'The effect of temperature on the proteome of recombinant *Pichia pastoris'*, *J Proteome Res*, 8. doi: 10.1021/pr8007623.

Dragosits, M., Stadlmann, J., Albiol, J., Baumann, K., Maurer, M., Gasser, B., Sauer, M., Altmann, F., Ferrer, P. and Mattanovich, D. (2009) 'The effect of temperature on the proteome of recombinant *Pichia pastoris*', *Journal of Proteome Research*, 8(3), pp. 1380–1392. doi: 10.1021/pr8007623.

Duetz, W. A., Ruedi, L., Hermann, R., O'Connor, K., Buchs, J. and Witholt, B. (2000) 'Methods for intense aeration, growth, storage, and replication of bacterial strains in microtiter plates', *Applied and Environmental Microbiology*, 66(6), pp. 2641–2646. doi: 10.1128/AEM.66.6.2641-2646.2000.

Edlund, S., Kaufman, J., Lessler, J., Douglas, J., Bromberg, M., Kaufman, Z., Bassal, R., Chodick, G., Marom, R., Shalev, V., Mesika, Y., Ram, R. and Leventhal, A. (2011) 'Comparing three basic models for seasonal influenza', *Epidemics*, 3(3–4), pp. 135–142. doi: 10.1016/j.epidem.2011.04.002.

Ellis, R. W., Rappuoli, R. and Ahmed, S. (2013) 'Technologies for making new vaccines', in *Vaccines*. Elsevier, pp. 1182–1199. doi: 10.1016/B978-1-4557-0090-5.00013-6.

EMA (2013) 'Fluenz Tetra® Assessment report', 44(June), pp. 1–91. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/002617/WC500158413.pdf.

FDA (2013) FLuBlok. *U. S. Food and Drug Administration*. Available at: http://www.fda.gov/BiologicsBloodVaccines/Vaccines/ApprovedProducts/ucm335836.htm (Accessed: 1 January 2014).

Fiers, W., De Filette, M., El Bakkouri, K., Schepens, B., Roose, K., Schotsaert, M., Birkett, A. and Saelens, X. (2009) 'M2e-based universal influenza A vaccine.', *Vaccine*, 27(45), pp. 6280–3. doi: 10.1016/j.vaccine.2009.07.007.

Follows, M., Hetherington, P. J., Dunnill, P. and Lilly, M. D. (1971) 'Release of enzymes from bakers' yeast by disruption in an industrial homogenizer', *Biotechnology and Bioengineering*, 13(4), pp. 549–560. doi: 10.1002/bit.260130408.

Freivalds, J., Dislers, A., Ose, V., Pumpens, P., Tars, K. and Kazaks, A. (2011) 'Highly efficient production of phosphorylated hepatitis B core particles in yeast *Pichia pastoris*', *Protein Expression and Purification*. Elsevier Inc., 75(2), pp. 218–224. doi: 10.1016/j.pep.2010.09.010.

Genzel, Y., Behrendt, I., König, S., Sann, H. and Reichl, U. (2004) 'Metabolism of MDCK cells during cell growth and influenza virus production in large-scale microcarrier culture', *Vaccine*, 22(17–18), pp. 2202–2208. doi: 10.1016/j.vaccine.2003.11.041.

Genzel, Y., Fischer, M. and Reichl, U. (2006) 'Serum-free influenza virus production avoiding washing steps and medium exchange in large-scale microcarrier culture', *Vaccine*, 24(16), pp. 3261–3272. doi: 10.1016/j.vaccine.2006.01.019.

Gerdil, C. (2003) 'The annual production cycle for influenza vaccine.', *Vaccine*, 21(16), pp. 1776–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12686093.

Glaxo Smith Kline (2005) *Boostrix*. Available at: http://us.gsk.com/products/assets/us_boostrix.pdf.

Gurramkonda, C., Zahid, M., Nemani, S. K., Adnan, A., Gudi, S. K., Khanna, N., Ebensen, T., Lunsdorf, H., Guzman, C. A. and Rinas, U. (2013) 'Purification of hepatitis B surface antigen virus-like particles from recombinant *Pichia pastoris* and in vivo analysis of their immunogenic properties', *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*. Elsevier B.V., 940, pp. 104–111. doi: 10.1016/j.jchromb.2013.09.030.

Haaheim, L. R. and Oxford, J. S. (2013) 'Basic influenza virology and immunology.', in *Pandemic influenza*. Wallingford: CABI, pp. 19–30. doi: 10.1079/9781845938567.0019.

Hartner, F. S. and Glieder, A. (2006) 'Regulation of methanol utilisation pathway genes in yeasts', 21, pp. 1–21. doi: 10.1186/1475-2859-5-39.

Heinen, P. P., Rijsewijk, F. A., de Boer-Luijtze, E. A. and Bianchi, A. T. J. (2002) 'Vaccination of pigs with a DNA construct expressing an influenza virus M2-nucleoprotein fusion protein exacerbates disease after challenge with influenza A virus.', *The Journal of general virology*, 83(Pt 8), pp. 1851–9. doi: 10.1099/0022-1317-83-8-1851.

Hemsworth, G. R., Henrissat, B., Davies, G. J. and Walton, P. H. (2014) 'Discovery and characterization of a new family of lytic polysaccharide monooxygenases', *Nat Chem Biol*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 10(2), pp. 122–126. Available at: http://dx.doi.org/10.1038/nchembio.1417.

Hewitt, C. J., Nebe-Von Caron, G., Axelsson, B., McFarlane, C. M. and Nienow, A. W. (2000) 'Studies related to the scale-up of high-cell-density *E. coli* fed-batch fermentations using multiparameter flow cytometry: Effect of a changing microenvironment with respect to glucose and dissolved oxygen concentration', *Biotechnology and Bioengineering*, 70(4), pp. 381–390. doi: 10.1002/1097-0290(20001120)70:4<381::AID-BIT3>3.0.CO;2-0.

Hink, M. A., Griep, R. A., Borst, J. W., Van Hoek, A., Eppink, M. H. M., Schots, A. and Visser, A. J. W. G. (2000) 'Structural dynamics of green fluorescent protein alone

and fused with a single chain Fv protein', *Journal of Biological Chemistry*, 275(23), pp. 17556–17560. doi: 10.1074/jbc.M001348200.

Holmes, K., Shepherd, D. A., Ashcroft, A. E., Whelan, M., Rowlands, D. J. and Stonehouse, N. J. (2015) 'Assembly pathway of hepatitis B core virus-like particles from genetically fused dimers', *Journal of Biological Chemistry*, 290(26), pp. 16238–16245. doi: 10.1074/jbc.M114.622035.

Hsieh, P. and Robbins, P. W. (1984) 'Regulation of asparagine-linked oligosaccharide processing. Oligosaccharide processing in Aedes albopictus mosquito cells.', *The Journal of Biological Chemistry*, 259(4), pp. 2375–82. Available at: http://www.ncbi.nlm.nih.gov/pubmed/6698972.

Huleatt, J. W., Nakaar, V., Desai, P., Huang, Y., Hewitt, D., Jacobs, A., Tang, J., McDonald, W., Song, L., Evans, R. K., Umlauf, S., Tussey, L. and Powell, T. J. (2008) 'Potent immunogenicity and efficacy of a universal influenza vaccine candidate comprising a recombinant fusion protein linking influenza M2e to the TLR5 ligand flagellin', *Vaccine*, 26(2), pp. 201–214. doi: 10.1016/j.vaccine.2007.10.062.

Hwang, S., Shao, Q., Williams, H., Hilty, C. and Gao, Y. Q. (2011) 'Methanol strengthens hydrogen bonds and weakens hydrophobic interactions in proteins--a combined molecular dynamics and NMR study.', *The journal of physical chemistry. B*, 115(20), pp. 6653–60. doi: 10.1021/jp111448a.

ICTV (2012) *International Committee on Taxonomy of Viruses*. Available at: http://www.ictvonline.org/virusTaxonomy.asp?version=2012 (Accessed: 1 January 2014).

Imperial College London (2000) *Immunology*. Available at: http://www.ch.ic.ac.uk/local/projects/sanderson/immunology.htm (Accessed: 1 January 2014).

Invitrogen Corporation (2002) 'Pichia Fermentation Process Guidelines Overview Overview , continued', pp. 1–11. Available at: https://tools.thermofisher.com/content/sfs/manuals/pichiaferm_prot.pdf.

IQur (2014) *Tandem Core*. Available at: http://www.iqur.com/index.php/tandem-core (Accessed: 1 January 2014).

Islam, R. S., Tisi, D., Levy, M. S. and Lye, G. J. (2008) 'Scale-up of *Escherichia coli* growth and recombinant protein expression conditions from microwell to laboratory and pilot scale based on matchedkLa', *Biotechnology and Bioengineering*, 99(5), pp. 1128–1139. doi: 10.1002/bit.21697.

Jegerlehner, A., Storni, T., Lipowsky, G., Schmid, M., Pumpens, P. and Bachmann, M. F. (2002) 'Regulation of IgG antibody responses by epitope density and CD21-mediated costimulation', *European Journal of Immunology*, 32(11), pp. 3305–3314. doi: 10.1002/1521-4141(200211)32:11<3305::AID-IMMU3305>3.0.CO;2-J.

Jenning, V., Lippacher, a and Gohla, S. H. (2002) 'Medium scale production of solid lipid nanoparticles (SLN) by high pressure homogenization.', *Journal of microencapsulation*, 19(1), pp. 1–10. doi: 10.1080/713817583.

Jiang, Z., Tong, G., Cai, B., Xu, Y. and Lou, J. (2011) 'Purification and immunogenicity study of human papillomavirus 58 virus-like particles expressed in *Pichia pastoris*', *Protein Expression and Purification*. Elsevier Inc., 80(2), pp. 203–210. doi: 10.1016/j.pep.2011.07.009.

Johnson, M. (2013) *Detergents: Triton X-100, Tween-20, and More*. doi: dx.doi.org/10.13070/mm.en.3.163.

Josefsberg, J. O. and Buckland, B. (2012) 'Vaccine process technology.', *Biotechnology and bioengineering*, 109(6), pp. 1443–60. doi: 10.1002/bit.24493.

Junker, B. H. (2004) 'Scale-up methodologies for *Escherichia coli* and yeast fermentation processes.', *Journal of Bioscience and Bioengineering*, 97(6), pp. 347–364. doi: 10.1263/jbb.97.347.

Kawai, N., Ikematsu, H., Hirotsu, N., Maeda, T., Kawashima, T., Tanaka, O., Yamauchi, S., Kawamura, K., Matsuura, S., Nishimura, M., Iwaki, N. and Kashiwagi, S. (2009) 'Clinical Effectiveness of Oseltamivir and Zanamivir for Treatment of Influenza A Virus Subtype H1N1 with the H274Y Mutation: A Japanese, Multicenter Study of the 2007–2008 and 2008–2009 Influenza Seasons', *Clinical Infectious Diseases*, 49(12), pp. 1828–1835. doi: 10.1086/648424.

Kazaks, A., Lu, I. N., Farinelle, S., Ramirez, A., Crescente, V., Blaha, B., Ogonah, O., Mukhopadhyay, T., de Obanos, M. P., Krimer, A., Akopjana, I., Bogans, J., Ose, V., Kirsteina, A., Kazaka, T., Stonehouse, N. J., Rowlands, D. J., Muller, C. P., Tars, K. and Rosenberg, W. M. (2017) 'Production and purification of chimeric HBc virus-like particles carrying influenza virus LAH domain as vaccine candidates', *BMC Biotechnology*, 17(1), pp. 1–11. doi: 10.1186/s12896-017-0396-8.

Kee, G. S., Pujar, N. S. and Titchener-Hooker, N. J. (2008) 'Study of Detergent-Mediated Liberation of Hepatitis B Virus-like Particles from S. cerevisiae Homogenate: Identifying a Framework for the Design of Future-Generation Lipoprotein Vaccine Processes', *Biotechnology Progress*. Wiley Subscription Services, Inc., A Wiley Company, 24(3), pp. 623–631. doi: 10.1021/bp070472i.

Kensy, F., Engelbrecht, C. and Büchs, J. (2009a) 'Scale-up from microtiter plate to laboratory fermenter: evaluation by online monitoring techniques of growth and protein expression in *Escherichia coli* and *Hansenula polymorpha fermentations*', *Microbial Cell Factories*, 8(1), p. 68. doi: 10.1186/1475-2859-8-68.

Kensy, F., Engelbrecht, C. and Büchs, J. (2009b) 'Scale-up from microtiter plate to laboratory fermenter: evaluation by online monitoring techniques of growth and protein expression in *Escherichia coli* and *Hansenula polymorpha* fermentations.', *Microbial cell factories*, 8, p. 68. doi: 10.1186/1475-2859-8-68.

Kim, E.-H., Lee, J.-H., Pascua, P. N. Q., Song, M.-S., Baek, Y.-H., Kwon, H.-I., Park, S.-J., Lim, G.-J., Decano, A., Chowdhury, M. Y., Seo, S.-K., Song, M. K., Kim, C.-J. and

Choi, Y.-K. (2013) 'Prokaryote-expressed M2e protein improves H9N2 influenza vaccine efficacy and protection against lethal influenza A virus in mice.', *Virology journal*, 10, p. 104. doi: 10.1186/1743-422X-10-104.

Kirk, T. V. and Szita, N. (2013) 'Oxygen transfer characteristics of miniaturized bioreactor systems', *Biotechnology and Bioengineering*, 110(4), pp. 1005–1019. doi: 10.1002/bit.24824.

Kleinig, A. R., Mansell, C. J., Nguyen, Q. D., Badalyan, A. and Middelberg, A. P. J. (1995) 'Influence of broth dilution on the disruption of *Escherichia coli*', *Biotechnology Techniques*, 9(10), pp. 759–762. doi: 10.1007/BF00159244.

Kleinig, A. R. and Middelberg, A. P. J. (1997) 'On the mechanism of microbial cell disruption in high pressure homogenisation', *Chemical Engineering Science*, 53(5), pp. 891–898. doi: 10.1016/S0009-2509(97)00414-4.

Klimek-Ochab, M., Brzezińska-Rodak, M., Zymańczyk-Duda, E., Lejczak, B. and Kafarski, P. (2011) 'Comparative study of fungal cell disruption-scope and limitations of the methods', *Folia Microbiologica*, 56(5), pp. 469–475. doi: 10.1007/s12223-011-0069-2.

Klöckner, W. and Büchs, J. (2012) 'Advances in shaking technologies', *Trends in Biotechnology*, 30(6), pp. 307–314. doi: 10.1016/j.tibtech.2012.03.001.

Krainer, F. W., Dietzsch, C., Hajek, T., Herwig, C., Spadiut, O. and Glieder, A. (2012) 'Recombinant protein expression in *Pichia pastoris* strains with an engineered methanol utilization pathway', *Microbial Cell Factories*, 11(1), p. 22. doi: 10.1186/1475-2859-11-22.

Krammer, F. and Palese, P. (2013) 'Influenza virus hemagglutinin stalk-based antibodies and vaccines', *Curr Opin Virol.*, (October), pp. 521–530. doi: doi:10.1016/j.coviro.2013.07.007.

Kreijtz, J. H. C. M., Fouchier, R. A. M. and Rimmelzwaan, G. F. (2011) 'Immune responses to influenza virus infection', *Virus Research*, 162(1–2), pp. 19–30. doi: 10.1016/j.virusres.2011.09.022.

Kuno, G. and Chang, G.-J. J. (2005) 'Biological transmission of arboviruses: reexamination of and new insights into components, mechanisms, and unique traits as well as their evolutionary trends.', *Clinical Microbiology Reviews*, 18(4), pp. 608–37. doi: 10.1128/CMR.18.4.608-637.2005.

Kuroda, K., Hauser, C., Rott, R., Klenk, H. D. and Doerfler, W. (1986) 'Expression of the influenza virus haemagglutinin in insect cells by a baculovirus vector.', *The EMBO journal*, 5(6), pp. 1359–65. Available at: http://www.ncbi.nlm.nih.gov/pubmed/3015601.

Layton, C. and Lenfestey, N. (2005a) 'Influenza Vaccine Economics', (208665).

Layton, C. and Lenfestey, N. (2005b) *Influenza Vaccine Manufacturing*.

Lee, Jae-Dong and Komagata, K. (1980) 'Taxonomic study of methanol assimilating yeasts', 158, pp. 133–158.

Lee, C. Y., Lee, S. J., Jung, K. H., Katoh, S. and Lee, E. K. (2003) 'High dissolved oxygen tension enhances heterologous protein expression by recombinant *Pichia pastoris*', *Process Biochemistry*, 38(8), pp. 1147–1154. doi: 10.1016/S0032-9592(02)00280-7.

Li, J., Tang, C., Shi, H. and Wu, M. (2011) 'Cloning and optimized expression of a neutral endoglucanase gene (ncel5A) from Volvariella volvacea WX32 in *Pichia pastoris*', *Journal of Bioscience and Bioengineering*, 111(5), pp. 537–540. doi: 10.1016/j.jbiosc.2011.01.002.

Li, Q., Aucamp, J. P., Tang, A., Chatel, A. and Hoare, M. (2012) 'Use of focused acoustics for cell disruption to provide ultra scale-down insights of microbial homogenization and its bioprocess impact-recovery of antibody fragments from rec *E. coli*', *Biotechnology and Bioengineering*, 109(8), pp. 2059–2069. doi: 10.1002/bit.24484.

Li, Z., Xiong, F., Lin, Q., D'Anjou, M., Daugulis, A. J., Yang, D. S. and Hew, C. L. (2001) 'Low-temperature increases the yield of biologically active herring antifreeze protein in *Pichia pastoris*', *Protein Expr Purif*, 21. doi: 10.1006/prep.2001.1395.

Liu, C., Eichelberger, M. C., Compans, R. W. and Air, G. M. (1995) 'Influenza type A virus neuraminidase does not play a role in viral entry, replication, assembly, or budding.', *Journal of Virology*, 69(2), pp. 1099–1106. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC188682/.

Looser, V., Bruhlmann, B., Bumbak, F., Stenger, C., Costa, M., Camattari, A., Fotiadis, D. and Kovar, K. (2014) 'Cultivation strategies to enhance productivity of *Pichia pastoris*: A review', *Biotechnology Advances*. Elsevier B.V., 33(6), pp. 1177–1193. doi: 10.1016/j.biotechadv.2015.05.008.

Lowen, A. C., Mubareka, S., Steel, J. and Palese, P. (2007) 'Influenza Virus Transmission Is Dependent on Relative Humidity and Temperature', *PLoS Pathogens*, 3(10), p. e151. doi: 10.1371/journal.ppat.0030151.

Luckow, V. A. and Summers, M. D. (1989) 'High level expression of nonfused foreign genes with Autographa californica nuclear polyhedrosis virus expression vectors.', *Virology*, 170(1), pp. 31–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/2497580.

Ludgate, L., Liu, K., Luckenbaugh, L., Streck, N., Eng, S., Voitenleitner, C., Delaney, W. E. and Hu, J. (2016) 'Cell-Free Hepatitis B Virus Capsid Assembly Dependent on the Core Protein C-Terminal Domain and Regulated by Phosphorylation', *Journal of Virology*. Edited by G. McFadden, 90(12), pp. 5830–5844. doi: 10.1128/JVI.00394-16.

Maccani, A., Landes, N., Stadlmayr, G., Maresch, D., Leitner, C., Maurer, M., Gasser, B., Ernst, W., Kunert, R. and Mattanovich, D. (2014) '*Pichia pastoris* secretes recombinant proteins less efficiently than Chinese hamster ovary cells but allows higher space-time yields for less complex proteins.', *Biotechnology journal*, 9(4), pp. 526–37. doi: 10.1002/biot.201300305.

Madigan, M. T., Martinko, J. M., Dunlap, P. V. and Clark, D. P. (2009) 'Essentials of Immunology', in *Brock Biology of Microorganisms*. San Francisco: Pearson, pp. 839–899.

Mahy, B. W. J. (1985) *Virology – a practical approach*. Oxford: IRL Press.

Manmohit Kalia, P. K. (2015) 'Pectin Methylesterases: A Review', *Journal of Bioprocessing & Biotechniques*, 5(5). doi: 10.4172/2155-9821.1000227.

Martin, P. M. V and Martin-granel, E. (2006) 'Term Epidemic', 12(6).

Merten, O. W., Hannoun, C., Manuguerra, J. C., Ventre, F. and Petres, S. (1996) 'Production of influenza virus in cell cultures for vaccine preparation.', *Advances in experimental medicine and biology*, 397, pp. 141–51. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8718593.

Middelberg, A. P. J. (1995) 'Process-scale disruption of microorganisms', *Biotechnology Advances*, 13(3), pp. 491–551. doi: 10.1016/0734-9750(95)02007-P.

Middelberg, A. P. J. (2000) 'Microbial Cell Disruption by High-Pressure Homogenization', *Downstream Processing of Proteins Methods and Protocols*, 9, pp. 11–22.

Mirro, R. (2011) *Introduction to* Pichia Pastoris *in a Stirred-Tank Fermentor*. Eppendorf application note No 253.

Miyajima, Y., Satoh, K., Umeda, Y. and Makimura, K. (2009) 'Quantitation of fungal DNA contamination in commercial Zymolyase and Lyticase used in the preparation of fungi', *Japanese Journal of Medical Mycology*, 50(4), pp. 259–262. doi: 10.3314/jjmm.50.259.

Mohnen, D. (2008) 'Pectin structure and biosynthesis', *Current Opinion in Plant Biology*, 11(3), pp. 266–277. doi: 10.1016/j.pbi.2008.03.006.

Monto, A. S. and Sellwood, C. (2013) 'History and epidemiological features of pandemic influenza.', in *Pandemic influenza*. Wallingford: CABI, pp. 40–48. doi: 10.1079/9781845938567.0040.

MSE (2017) *Soniprep 150 Plus*. Available at: http://www.mseuk.co.uk/Products/UltrasonicDisintegrators/Soniprep_150_Plus/Default.aspx (Accessed: 1 January 2017).

Nakagawa, T., Yamada, K., Fujimura, S., Ito, T., Miyaji, T. and Tomizuka, N. (2005) 'Pectin utilization by the methylotrophic yeast Pichia methanolica', *Microbiology*, 151(6), pp. 2047–2052. doi: 10.1099/mic.0.27895-0.

Nakayama, T., Kashiwagi, Y., Kawashima, H., Kumagai, T., Ishii, K. J. and Ihara, T. (2012) 'Alum-adjuvanted H5N1 whole virion inactivated vaccine (WIV) enhanced inflammatory cytokine productions', *Vaccine*, 30(26), pp. 3885–3890. doi: 10.1016/j.vaccine.2012.04.004.

National Research Council Canada (2010) *National Research Council Canada*. Available at: http://www.nrc-cnrc.gc.ca/eng/dimensions/issue3/adjuvants.html (Accessed: 1 January 2014).

Nayak, D. P., Lehmann, S. and Reichl, U. (2005) 'Downstream processing of MDCK cell-derived equine influenza virus', *Journal of Chromatography B*, 823(2), pp. 75–81. doi: 10.1016/j.jchromb.2005.05.022.

Nerome, K., Kumihashi, H., Nerome, R., Hiromoto, Y., Yokota, Y., Ueda, R., Omoe, K. and Chiba, M. (1999) 'Evaluation of immune responses to inactivated influenza vaccines prepared in embryonated chicken eggs and MDCK cells in a mouse model.', *Developments in biological standardization*, 98, pp. 53-63–4. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10494959.

Neumann, G., Ozawa, M. and Kawaoka, Y. (2012) 'Reverse Genetics of Influenza Viruses', in, pp. 193–206. doi: 10.1007/978-1-61779-621-0_12.

Nicholson, K. G., Wood, J. M. and Zambon, M. (2003) 'Influenza', *The Lancet*, 362(9397), pp. 1733–1745. doi: 10.1016/S0140-6736(03)14854-4.

Nims, R. W. (2006) 'Detection of adventitious viruses in biologicals--a rare occurrence.', *Developments in biologicals*, 123, pp. 153-64-97. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16566443.

Novagen (2004) 'YeastBusterTM Protein Extraction Reagent', *User Protocol TB316 Rev. B 0804*, pp. 1–4. Available at: https://www.merckmillipore.com/GB/en/product/YeastBuster^{TM}-Protein-Extraction-Reagent,EMD_BIO-71186?ReferrerURL=https%3A%2F%2Fwww.google.co.uk%2F&bd=1#documentation.

Omstead, D. R. (1990) 'Compositional gas analyzer-based measurements', in *Computer control of fermentation processes*. Boca Raton, Florida: CRC Press, Inc., pp. 112–113.

Orman, M. A., Çelık, P. and Ozdamar, T. H. (2009) 'The influence of carbon sources on recombinant-human- growth-hormone production by *Pichia pastoris* is dependent on phenotype: a comparison of MutS and Mut+ strains', *Biotechnol. Appl. Biochem*, 52, pp. 245–255. doi: 10.1042/BA20080057.

Oxford, J. S., Corcoran, T., Knott, R., Bates, J., Bartolomei, O., Major, D., Newman, R. W., Yates, P., Robertson, J., Webster, R. G. and Schild, G. C. (1987) 'Serological

studies with influenza A(H1N1) viruses cultivated in eggs or in a canine kidney cell line (MDCK)', *Bulletin of the World Health Organization*, 65(2), pp. 181–187. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2490845/.

Oxford, J. S., Newman, R., Corcoran, T., Bootman, J., Major, D., Yates, P., Robertson, J. and Schild, G. C. (1991) 'Direct Isolation in Eggs of Influenza A (H1N1) and B Viruses with Haemagglutinins of Different Antigenic and Amino Acid Composition', *Journal of General Virology*, 72(1), pp. 185–189. doi: 10.1099/0022-1317-72-1-185.

Palache, A. M., Brands, R. and van Scharrenburg, G. J. (1997) 'Immunogenicity and reactogenicity of influenza subunit vaccines produced in MDCK cells or fertilized chicken eggs.', *The Journal of infectious diseases*, 176 Suppl, pp. S20-3. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9240689.

Palache, A. M., Scheepers, H. S., de Regt, V., van Ewijk, P., Baljet, M., Brands, R. and van Scharrenburg, G. J. (1999) 'Safety, reactogenicity and immunogenicity of Madin Darby Canine Kidney cell-derived inactivated influenza subunit vaccine. A meta-analysis of clinical studies.', *Developments in biological standardization*, 98, pp. 115-25–4. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10494964.

Panavas, T., Sanders, C. and Butt, T. R. (2009) 'SUMO Fusion Technology for Enhanced Protein Production in Prokaryotic and Eukaryotic Expression Systems', in, pp. 303–317. doi: 10.1007/978-1-59745-566-4_20.

Pang, I. K. and Iwasaki, A. (2011) 'Inflammasomes as mediators of immunity against influenza virus', *Trends in Immunology*, 32(1), pp. 34–41. doi: 10.1016/j.it.2010.11.004.

Percheson, P. B., Trépanier, P., Dugré, R. and Mabrouk, T. (1999) 'A Phase I, randomized controlled clinical trial to study the reactogenicity and immunogenicity of a new split influenza vaccine derived from a non-tumorigenic cell line.', *Developments in biological standardization*, 98, pp. 127-32–4. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10494965.

Peyret, H., Gehin, A., Thuenemann, E. C., Blond, D., El Turabi, A., Beales, L., Clarke, D., Gilbert, R. J. C., Fry, E. E., Stuart, D. I., Holmes, K., Stonehouse, N. J., Whelan, M., Rosenberg, W., Lomonossoff, G. P. and Rowlands, D. J. (2015) 'Tandem fusion of hepatitis B core antigen allows assembly of virus-like particles in bacteria and plants with enhanced capacity to accommodate foreign proteins', *PLoS ONE*, 10(4), pp. 1–20. doi: 10.1371/journal.pone.0120751.

Pichlmair, A., Schulz, O., Tan, C. P., Naslund, T. I., Liljestrom, P., Weber, F. and Reis e Sousa, C. (2006) 'RIG-I-Mediated Antiviral Responses to Single-Stranded RNA Bearing 5'-Phosphates', *Science*, 314(5801), pp. 997–1001. doi: 10.1126/science.1132998.

Pinto, L. H., Holsinger, L. J. and Lamb, R. A. (1992) 'Influenza virus M2 protein has ion channel activity.', *Cell*, 69(3), pp. 517–28. Available at: http://www.ncbi.nlm.nih.gov/pubmed/1374685.

Plotch, S. J., Bouloy, M. and Krug, R. M. (1979) 'Transfer of 5'-terminal cap of globin mRNA to influenza viral complementary RNA during transcription in vitro.', *Proceedings of the National Academy of Sciences of the United States of America*, 76(4), pp. 1618–22. Available at: http://www.ncbi.nlm.nih.gov/pubmed/287003.

Protein Sciences Corporation (2013) *Flublok*. Available at: http://www.flublok.com (Accessed: 1 January 2014).

Robertson, J. S., Bootman, J. S., Newman, R., Oxford, J. S., Daniels, R. S., Webster, R. G. and Schild, G. C. (1987) 'Structural changes in the haemagglutinin which accompany egg adaptation of an influenza A(H1N1) virus.', *Virology*, 160(1), pp. 31–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/3629978.

Roland, D. (2014) *From chicken egg to syringe: how a flu vaccine is made*.

Roldao, A., Mellado, M. C., Castilho, L. R., Carrondo, M. J. and Alves, P. M. (2010) 'Virus-like particles in vaccine development', *Expert Review of Vaccines*, 9(10), pp. 1149–1176. doi: 10.1586/erv.10.115.

Roldão, A., Mellado, M. C. M., Castilho, L. R., Carrondo, M. J. and Alves, P. M. (2010) 'Virus-like particles in vaccine development', *Expert Review of Vaccines*, 9(10), pp. 1149–1176. doi: 10.1586/erv.10.115.

Roose, K., Baets, S. De, Schepens, B. and Saelens, X. (2013) 'Hepatitis B core–based virus–like particles to present heterologous epitopes', *Expert Review of Vaccines*, 12(2), pp. 183–198. doi: 10.1586/erv.12.150.

Salazar, O. and Asenjo, J. A. (2007) 'Enzymatic lysis of microbial cells', *Biotechnology Letters*, 29(7), pp. 985–994. doi: 10.1007/s10529-007-9345-2.

Samji, T. (2009) 'Influenza A: Understanding the Viral Life Cycle', *The Yale Journal of Biology and Medicine*. YJBM, 82(4), pp. 153–159. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2794490/.

Schmidt, F. R. (2005) 'Optimization and scale up of industrial fermentation processes', *Applied Microbiology and Biotechnology*, 68(4), pp. 425–435. doi: 10.1007/s00253-005-0003-0.

Shi, L., Sanyal, G., Ni, A., Luo, Z., Doshna, S., Wang, B., Graham, T. L., Wang, N. and Volkin, D. B. (2005) 'Stabilization of human papillomavirus virus-like particles by non-ionic surfactants', *Journal of Pharmaceutical Sciences*, 94(7), pp. 1538–1551. doi: 10.1002/jps.20377.

Sigma-Aldrich (2017) *Lyticase from Arthrobacter luteus*. Available at: https://www.sigmaaldrich.com/catalog/product/sigma/l2524?lang=en&region=GB (Accessed: 4 March 2017).

Skehel, J. J. and Wiley, D. C. (2000) 'Receptor Binding and Membrane Fusion in Virus Entry: The Influenza Hemagglutinin', *Annual Review of Biochemistry*, 69(1), pp. 531–569. doi: 10.1146/annurev.biochem.69.1.531.

Slingsby, F. and Dewar, S. (2015) *Use of the ambr ® 250 in combination with high-throughput design and analysis tools for rapid , scalable USP development*, *Upstream Process Development*. Available at: https://www.fujifilmdiosynth.com/assets/Ambr250_White_paper_Jan15_final.pdf

Sorrell, E., Schrauwen, E., Linster, M., De Graaf, M., Herfst, S. and Fouchier, R. (2011) 'Predicting "airborne" influenza viruses: (trans-) mission impossible?', *Current Opinion in Virology*, 1(6), pp. 635–642. doi: 10.1016/j.coviro.2011.07.003.

Strods, A., Ose, V., Bogans, J., Cielens, I., Kalnins, G., Radovica, I., Kazaks, A., Pumpens, P. and Renhofa, R. (2015) 'Preparation by alkaline treatment and detailed characterisation of empty hepatitis B virus core particles for vaccine and gene therapy applications', *Scientific Reports*. Nature Publishing Group, 5(May), p. 11639. doi: 10.1038/srep11639.

Summers, M., Smith, G. E. and Station., T. A. E. (1987) 'A manual of methods for baculovirus vectors and insect cell culture procedures'. College Station, Tex.: Texas Agricultural Experiment Station (Bulletin / Texas Agricultural Experiment Station ;no. 1555), p. 56 p. Available at: file://catalog.hathitrust.org/Record/100008078.

Szymczakiewicz-Multanowska, A., Groth, N., Bugarini, R., Lattanzi, M., Casula, D., Hilbert, A., Tsai, T. and Podda, A. (2009) 'Safety and Immunogenicity of a Novel Influenza Subunit Vaccine Produced in Mammalian Cell Culture', *The Journal of Infectious Diseases*, 200(6), pp. 841–848. doi: 10.1086/605505.

Talbot, H. K., Nian, H., Zhu, Y., Chen, Q., Williams, J. V. and Griffin, M. R. (2015) 'Clinical Effectiveness of Split-Virion Versus Subunit Trivalent Influenza Vaccines in Older Adults', *Clinical Infectious Diseases*, 60(8), pp. 1170–1175. doi: 10.1093/cid/civ019.

Tong, S., Li, Y., Rivailler, P., Conrardy, C., Castillo, D. A. A., Chen, L.-M., Recuenco, S., Ellison, J. A., Davis, C. T., York, I. A., Turmelle, A. S., Moran, D., Rogers, S., Shi, M., Tao, Y., Weil, M. R., Tang, K., Rowe, L. A., Sammons, S., Xu, X., Frace, M., Lindblade, K. A., Cox, N. J., Anderson, L. J., Rupprecht, C. E. and Donis, R. O. (2012) 'A distinct lineage of influenza A virus from bats', *Proceedings of the National Academy of Sciences*, 109(11), pp. 4269–4274. doi: 10.1073/pnas.1116200109.

Tree, J. A., Richardson, C., Fooks, A. R., Clegg, J. C. and Looby, D. (2001) 'Comparison of large-scale mammalian cell culture systems with egg culture for the production of influenza virus A vaccine strains', *Vaccine*, 19(25–26), pp. 3444–3450. doi: 10.1016/S0264-410X(01)00053-6.

Tschopp, J. F., Brust, P. F., Cregg, J. M., Stillman, C. A. and Gingeras, T. R. (1987) 'Expression of the lacZ gene from two methanol-regulated promoters in *Pichia*

pastoris.', *Nucleic Acids Research*, 15(9), pp. 3859–3876. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC340787/.

Van-Tam, J. and Lim, W. (2013) 'Pharmaceutical interventions.', in *Pandemic influenza*. Wallingford: CABI, pp. 122–138. doi: 10.1079/9781845938567.0122.

Van-Tam, J. and Sellwood, C. (2013) 'Epidemiology and clinical features of interpandemic influenza.', in *Pandemic influenza*. Wallingford: CABI, pp. 1–8. doi: 10.1079/9781845938567.0001.

Verma, R., Boleti, E. and George, A. J. (1998) 'Antibody engineering: comparison of bacterial, yeast, insect and mammalian expression systems.', *Journal of immunological methods*, 216(1–2), pp. 165–81. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9760222.

Virology Blog (2009) *Structure of influenza virus*. Available at: http://www.virology.ws/2009/04/30/structure-of-influenza-virus/ (Accessed: 20 August 2017).

Ward, D. P., Cárdenas-fernández, M., Hewitson, P., Ignatova, S. and Lye, G. J. (2015) 'Centrifugal partition chromatography in a biorefinery context : Separation of monosaccharides from hydrolysed sugar beet pulp ☹', *Journal of Chromatography A*. Elsevier B.V., 1411, pp. 84–91. doi: 10.1016/j.chroma.2015.08.006.

Ward, D. P., Hewitson, P., Cárdenas-fernández, M., Hamley-bennett, C., Díaz-rodríguez, A., Douillet, N., Adams, J. P., Leak, D. J., Ignatova, S. and Lye, G. J. (2017) 'Centrifugal partition chromatography in a biorefinery context : Optimisation and scale-up of monosaccharide fractionation from hydrolysed sugar beet pulp ☹', 1497, pp. 56–63.

Weinacker, D., Rabert, C., Zepeda, A. B., Figueroa, C. A., Pessoa, A. and Farías, J. G. (2013) 'Applications of recombinant *Pichia pastoris* in the healthcare industry', *Brazilian Journal of Microbiology*, 44(4), pp. 1043–1048. doi: 10.1590/S1517-83822013000400004.

Wenger, M. D., DePhillips, P., Bracewell, D. G., University College London and Merck & Co. (2008) 'A microscale yeast cell disruption technique for integrated process development strategies', *Biotechnology Progress*, 24(3), pp. 606–614. doi: 10.1021/bp070359s.

WHO (2009) *Pandemic inluenza vaccine manufacturing process and timetable*. Available at: http://www.who.int/csr/disease/swineflu/notes/h1n1_vaccine_20090806/en/index.html (Accessed: 1 January 2017).

Wolff, M. W. and Reichl, U. (2008) 'Downstream Processing: From Egg to Cell Culture-Derived Influenza Virus Particles', *Chemical Engineering & Technology*, 31(6), pp. 846–857. doi: 10.1002/ceat.200800118.

Wood, J. M., Oxford, J. S., Dunleavy, U., Newman, R. W., Major, D. and Robertson, J. S. (1989) 'Influenza A (H1N1) vaccine efficacy in animal models is influenced by two amino acid substitutions in the hemagglutinin molecule.', *Virology*, 171(1), pp. 214–21. Available at: http://www.ncbi.nlm.nih.gov/pubmed/2741341.

Wu, C.-Y., Yeh, Y.-C., Yang, Y.-C., Chou, C., Liu, M.-T., Wu, H.-S., Chan, J.-T. and Hsiao, P.-W. (2010) 'Mammalian Expression of Virus-Like Particles for Advanced Mimicry of Authentic Influenza Virus', *PLoS ONE*. Edited by C. Hauser, 5(3), p. e9784. doi: 10.1371/journal.pone.0009784.

Wu, W. W., Sun, Y.-H. B. and Pante, N. (2007) 'Nuclear import of influenza A viral ribonucleoprotein complexes is mediated by two nuclear localization sequences on viral nucleoprotein', *Virology Journal*, 4(1), p. 49. doi: 10.1186/1743-422X-4-49.

Xie, J., Zhou, Q., Du, P., Gan, R. and Ye, Q. (2005) 'Use of different carbon sources in cultivation of recombinant *Pichia pastoris* for angiostatin production', *Enzyme and Microbial Technology*, 36(2–3), pp. 210–216. doi: 10.1016/j.enzmictec.2004.06.010.

Zhan, D., Janssen, P. and Mort, A. J. (1998) 'Scarcity or complete lack of single rhamnose residues interspersed within the homogalacturonan regions of citrus pectin', *Carbohydrate Research*, 308(3–4), pp. 373–380. doi: 10.1016/S0008-6215(98)00096-2.

Zhong, Y., Yang, L., Guo, Y., Fang, F., Wang, D., Li, R., Jiang, M., Kang, W., Ma, J., Sun, J. and Xiao, W. (2014) 'High-temperature cultivation of recombinant *Pichia pastoris*

increases endoplasmic reticulum stress and decreases production of human interleukin-10', *Microbial Cell Factories*, 13(1), p. 163. doi: 10.1186/s12934-014-0163-7.

Zlotnick, A. and Mukhopadhyay, S. (2011) 'Virus assembly, allostery and antivirals', *Trends in Microbiology*. Elsevier Ltd, 19(1), pp. 14–23. doi: 10.1016/j.tim.2010.11.003.

Chapter 10 Appendix

10.1 Supplementary data (Chapter 2)

10.1.1 Media components

Table 10-1 Fermentation media components

10X YNB (Yeast Nitrogen Base with ammonium sulphate and without amino acids)	
Yeast Nitrogen Base (YNB) with ammonium sulphate and without amino acids	69 g
RO water	500 ml
500X B (0.02% Biotin)	
Biotin	20 mg
0.05 M NaOH	100 ml
10X GY (10% Glycerol)	
Glycerol	100 ml
RO water	900 ml
BMGY Medium (Buffered Glycerol-complex medium)	
Yeast Extract	10 g/l
Peptone	20 g/l
1M potassium phosphate, pH 6.0	100 ml/l
10X YNB	100 ml/l
500X B	2 ml/l
10X GY	100 ml/l
PTM_1 Trace Salts	
Cupric sulphate 5 H_2O	6.0 g/l
Sodium iodide	0.08 g/l
Manganese sulphate 7 H_2O	3.0 g/l
Sodium molybdate 2 H_2O	0.2 g/l
Boric acid	0.02 g/l
Cobalt chloride	0.5 g/l
Zinc chloride	20.0 g/l
Ferrous sulphate 7 H_2O	65.0 g/l
Biotin	0.2 g/l
Sulphuric acid	5.0 ml/l
Fermentation Basal Salts	
Phosphoric acid, 85 %	26.7 ml/l
Calcium sulphate	0.93 g/l
Potassium sulphate	18.2 g/l
Magnesium sulphate 7 H_2O	14.9 g/l
Potassium hydroxide	4.13 g/l
Glycerol	40.0 g/l
Antifoam (PPG2000 Foam Control Agent)	1 ml/l

10.1.2 Cell bank data

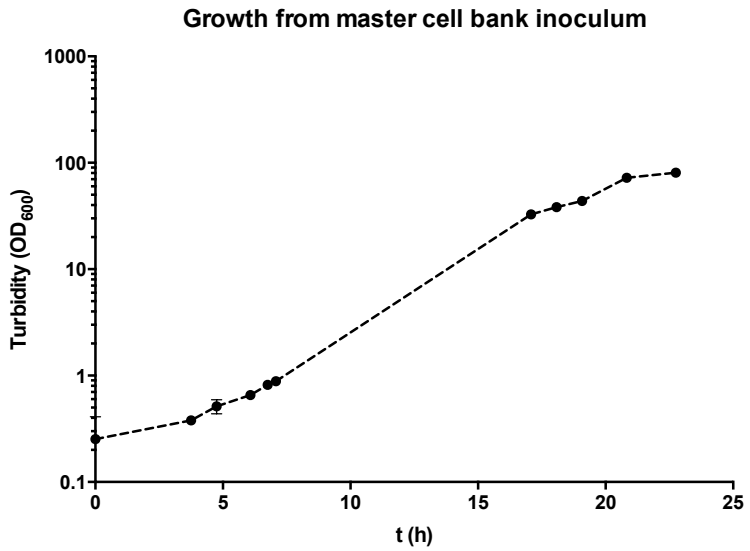

Figure 10.1 Growth profile of master cell bank *Pichia pastoris* KM71H expressing tHBc-k1,k1.

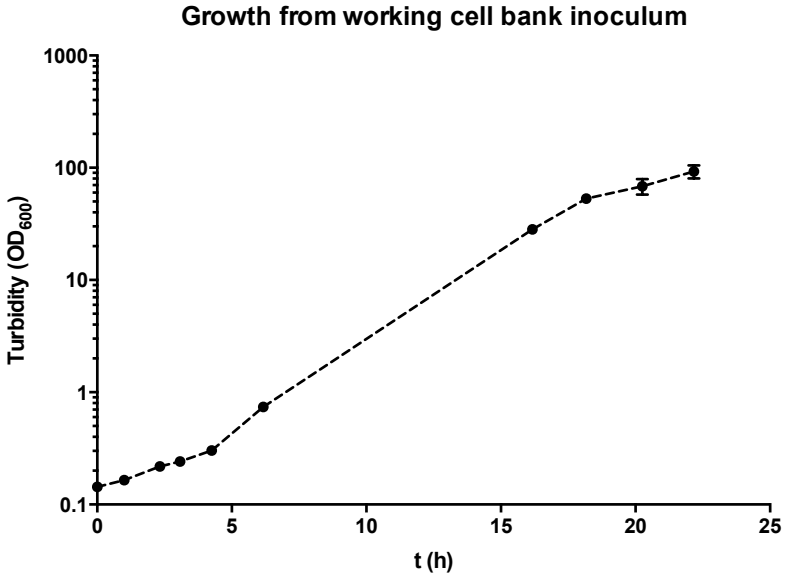

Figure 10.2 Growth profile of working cell bank *Pichia pastoris* KM71H expressing tHBc-k1,k1.

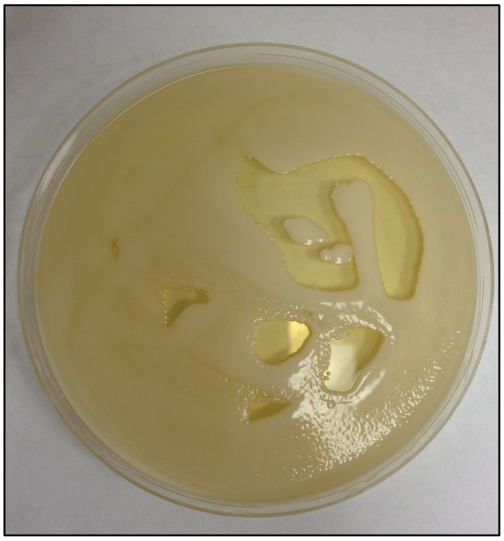

Figure 10.3 Non-contaminated agar culture of working cell bank *Pichia pastoris* X33 expressing tHBc-GFP,e at 28°C, 1 week.

10.2 Supplementary data (Chapter 3)

10.2.1 Fermentation graphs

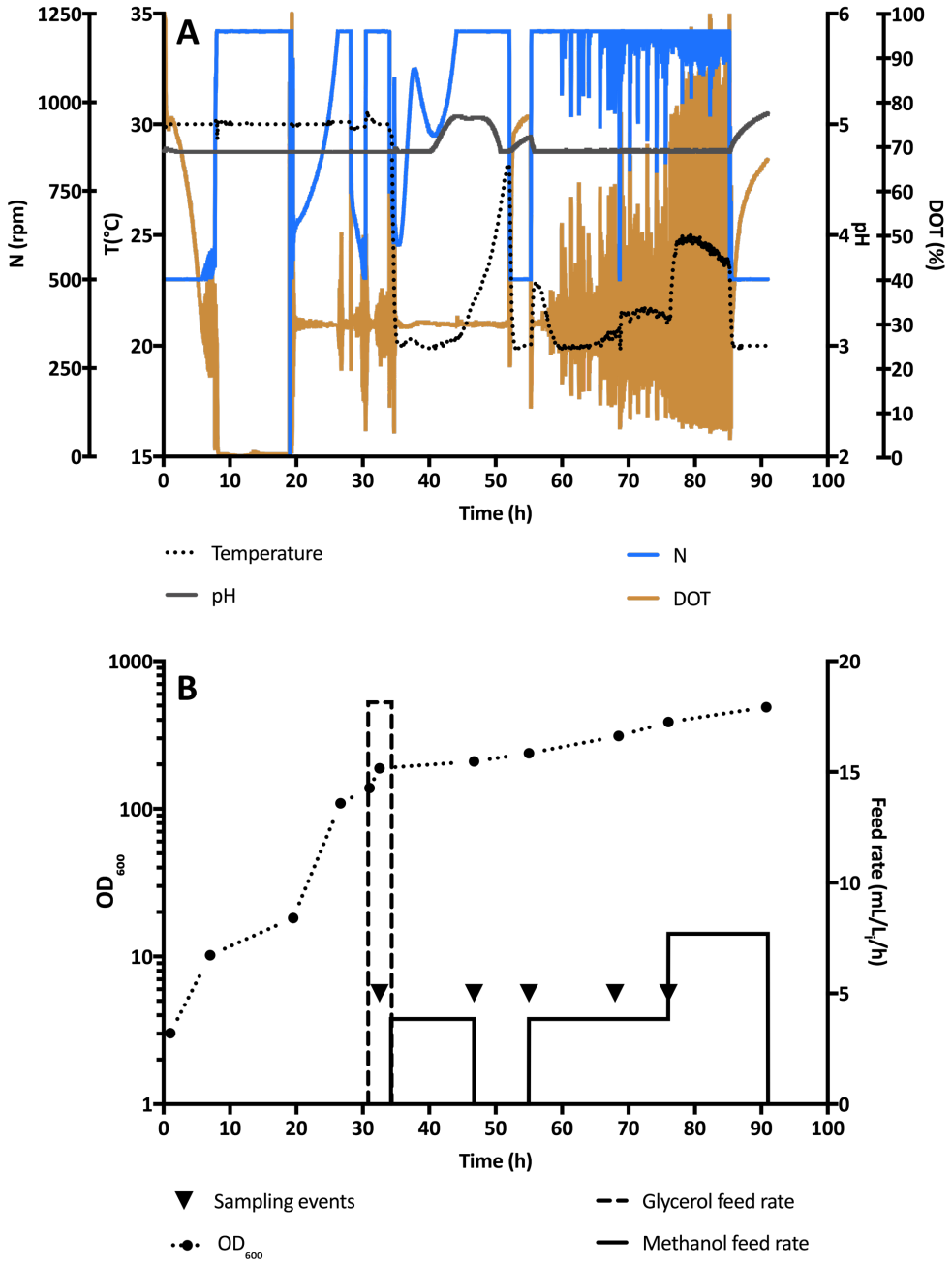

Figure 10.4 Low-temperature induction fermentation data of *P. pastoris* Mut$^+$ expressing tHBc-GFP,e: (A) Online data including agitation rate, temperature, pH and dissolved oxygen tension level. (B) Offline data including OD_{600}, sampling events and carbon source feed rates.

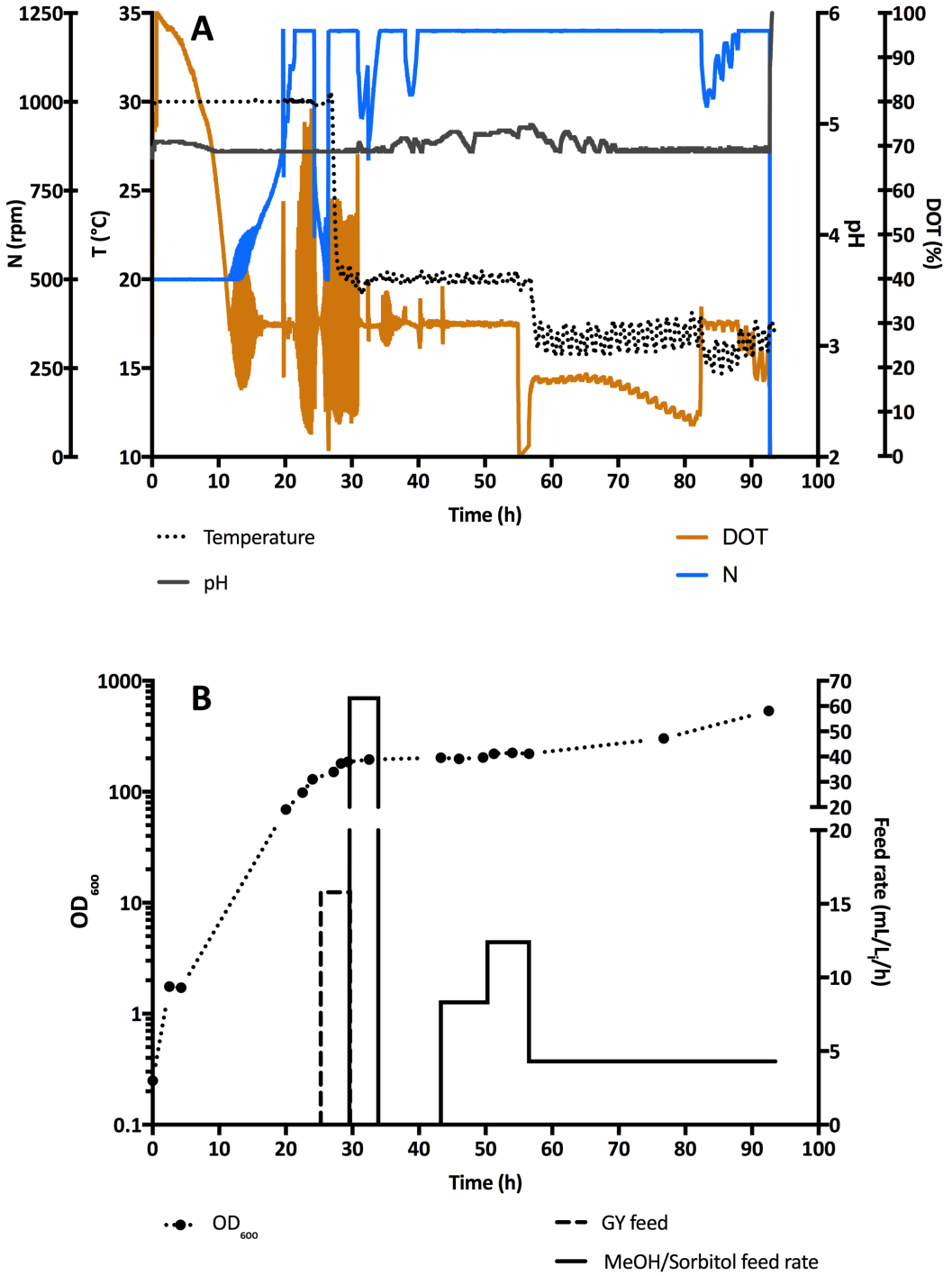

Figure 10.5 Mixed feeding induction fermentation data of *P.pastoris* Mut$^+$ expressing tHBc-GFP,e: (A) Online data including agitation rate, temperature, pH and dissolved oxygen tension level. (B) Offline data including OD_{600} and carbon source feed rates.

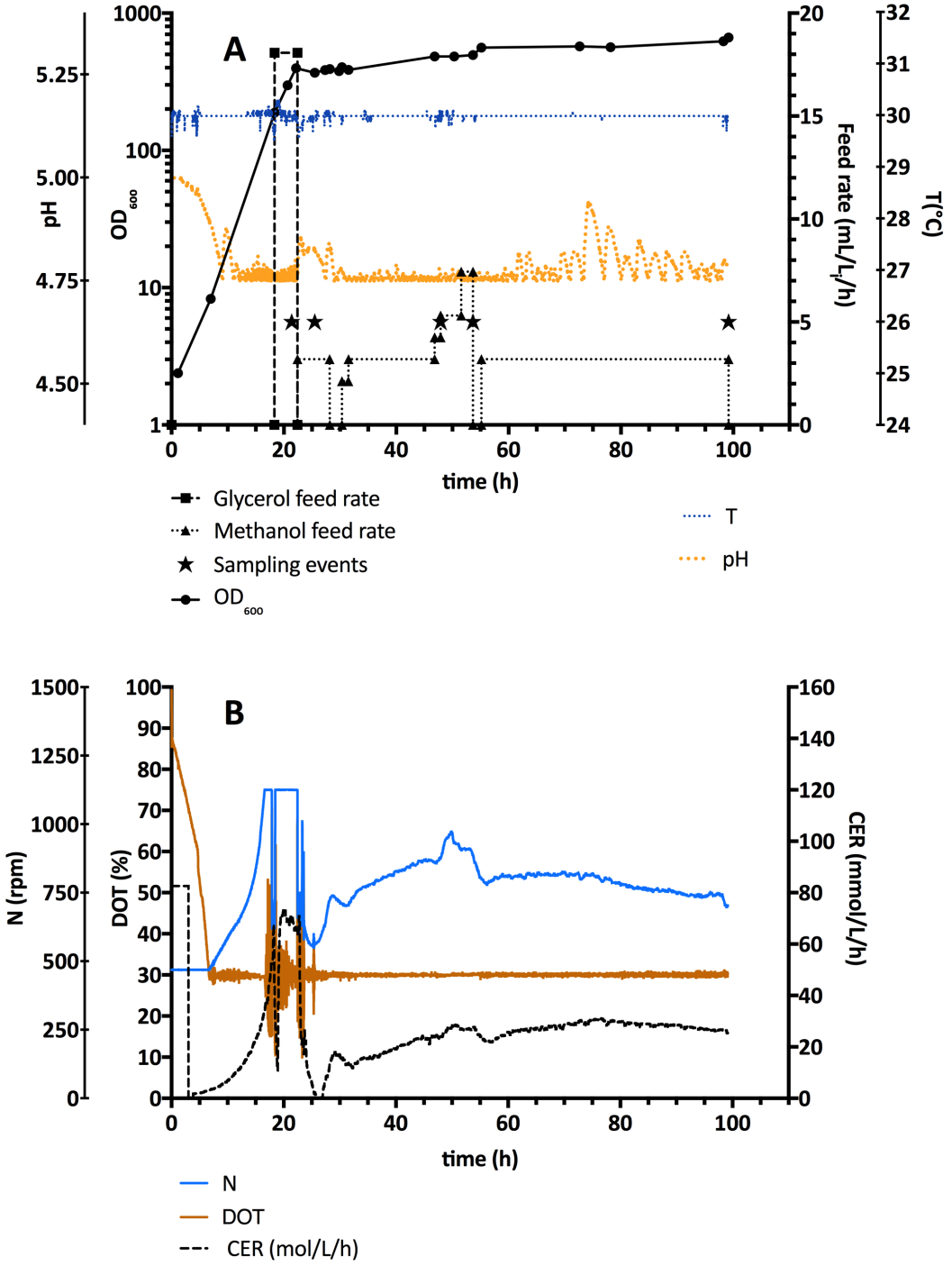

Figure 10.6 Fermentation data of *P. pastoris* Mut^S expressing tHBc-k1,k1: (A) Offline data including OD_{600}, sampling events, carbon source feed rates, temperature and pH (B) Online data including agitation rate, dissolved oxygen tension level and carbon evolution rate.

10.3 Supplementary data (Chapter 4)

10.3.1 Chemical lysis with YeastBuster™

YeastBuster™ is a protein extraction reagent produced by Novagen® with a proprietary mix of mild detergents, protein stabilisation buffer reducing agent called tris(hydroxypropyl)phosphine (THP). It has reportedly been used to perform extractions of proteins from *Pichia pastoris* through 'gentle conditions' (Novagen, 2004).

As this agent has previously been used to disrupt *Pichia pastoris*, it was hypothesised that YeastBuster™ could be effective in extracting soluble core-positive material from harvested cells.

The effectiveness of YeastBuster™ as a disruption agent was assessed by incubating harvested material (3.3.1) at various cell densities in (1) YeastBuster™ (2.4.8) and, for comparative purposes, (2) lyticase buffer (2.4.6). Note that the lyticase mix also contained 0.1% Triton-X100 prior to incubation.

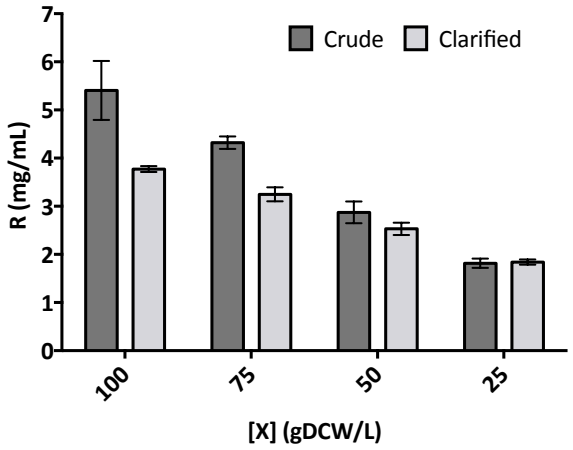

Figure 10.7 Total protein concentration of crude and clarified lysates of cells incubated in 1mL Yeastbuster™ (section 2.4.8) at various biomass concentrations ([X]). Averages and standard deviations are derived from triplicate measurements on duplicate samples for both crude and clarified lysates. Total protein quantification was determined using the Nanodrop method (section 2.5.5).

Figure 10.7 shows total protein concentrations of both crude and clarified lysates of cells incubated with Yeastbuster™ at various biomass concentrations. The data shows that Yeastbuster™ is capable of disrupting cells and that the total amount of total- and soluble protein released by Yeastbuster™ increases as the biomass concentration increases. However, the proportion of soluble protein to total protein measured decreases as the concentration of biomass increases.

The data shows that Yeastbuster™ is suitable for disrupting *P. pastoris*. However, the aim of using this agent is to recover soluble core-positive material. This was assessed using qualitative product analysis, which shown in Figure 10.8 and was performed on crude and clarified lysates using dot blot analysis as described in sections 2.5.10 and 2.5.11.

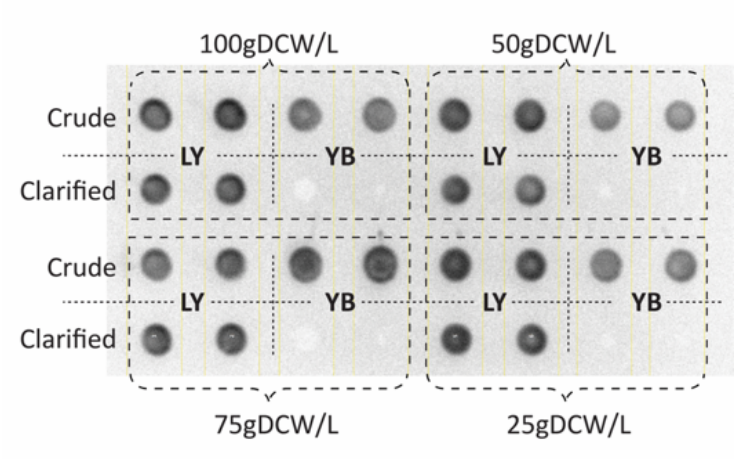

Figure 10.8 Qualitative analysis of crude and clarified lysates following lyticase (LY) (2000U/mL) treatment (2.4.6) and Yeastbuster™ (YB) treatments (2.4.8) at various concentrations of biomass at a total lysis volume of 1mL. Measurements were performed in duplicate. Reference material (hepatitis B core antigen) was used as a positive control (not shown).

Figure 10.8 shows no detectable recovery of soluble core-positive material when Yeastbuster™ is used. This demonstrates that, despite being capable of disrupting *Pichia pastoris*, certain unknown characteristics of Yeastbuster™ cause core-positive material to become insoluble. The use of Yeastbuster™ as an agent for cell disruption in future experiments was, therefore, not recommended.

10.3.2 Sonication

Cells can be disrupted through intense localised shockwaves created by ultrasonic waves. These waves can be produced by sonication devices such as the SoniPrep 150 (MSE, 2017). This method of disruption has been used for the disruption for a variety of microbial expression systems (Chakrabarti *et al.*, 2001; Hemsworth *et al.*, 2014), including *P. pastoris.* Section 2.4.4 shows how the SoniPrep 150 was used throughout this thesis. It was therefore hypothesised that this method could be a

suitable extraction method for the recovery of soluble core-positive material from harvested cells.

Total cell disruption through sonication of harvested material (3.2.1) was compared to material disrupted through lyticase as described in section 2.4.6. This comparison is shown in Figure 10.9.

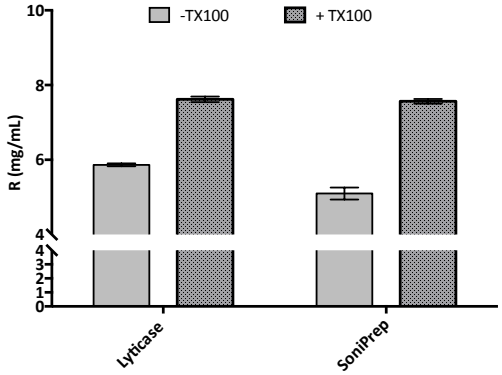

Figure 10.9 Comparison of lyticase-mediated cell disruption versus SoniPrep-mediated cell disruption with and without the addition of Triton-X100 after cell disruption. Triplicate measurements were performed on clarified lysate using the Nanodrop method (section 2.5.5).

It seems that total levels of cell disruption achieved through sonication are similar to those achieved through incubation with lyticase. However, section 3.2.2 had also shown that the implementation of sonication resulted in a loss of soluble core-positive material. It has been speculated that this could be due to heating effects. Furthermore, this method of cell disruption relies heavily on manual operation, presents an increased risk of sample cross-contamination and is not suitable for automation. Therefore, sonication with the SoniPrep 150 was not considered an ideal scale-down cell disruption platform.

10.4 Supplementary data (Chapter 5)

10.4.1 Theory of the effect of working volume on growth and expression

A change in working volume (V_w) leads to a change in the oxygen mass transfer coefficient, k_La, as shown in the following correlation (Doran, 1996):

$$\mathbf{k_L a = f(V_W)} \qquad \textbf{Equation 10-1}$$

In turn, the oxygen mass transfer coefficient affects the Oxygen Transfer Rate (OTR), which describes the amount of oxygen made available per unit of volume, per unit of time. This correlation is shown in the following equation (Klöckner and Büchs, 2012; Kirk and Szita, 2013):

$$\mathbf{OTR = k_L a(C^* - C_L)} \qquad \textbf{Equation 10-2}$$

Where:

- OTR represents the Oxygen Transfer Rate (mmol/L/h)
- C^* represents maximum concentration of dissolved oxygen at equilibrium (mmol/L)
- $\mathrm{C_L}$ represents the observed concentration of oxygen in the culture (mmoL/L)

The OTR is important as it can be used to define whether a culture is being supplied with sufficient oxygen, through comparison with the Oxygen Uptake Rate (OUR). This is because a for a culture to grow the OTR must be equal to or greater than the OUR.

The OUR describes the rate of oxygen consumption by a culture, which in turn is dependent on the specific growth rate as shown in the following equation:

$$\mathbf{OUR} = \frac{\mu}{\mathbf{Y_{X/O_2}}}[\mathbf{X}] + \mathbf{m}[\mathbf{X}] \qquad \textbf{Equation 10-3}$$

Where:

- OUR represents the Oxygen Uptake Rate (mmol/L/h)
- $\mathrm{Y_{X/O_2}}$ represents the yield of biomass over oxygen
- m represents a maintenance coefficient and can be determined experimentally

The yield of biomass over oxygen can be calculated using Equation 10-4. This equation is a representation of the metabolism of a culture.

$$\mathrm{C_wH_xO_yN_z} + \mathrm{a} \cdot \mathrm{O_2} + \mathrm{b} \cdot \mathrm{H_gO_hN_i} \rightarrow \mathrm{c} \cdot \mathrm{CH_\alpha O_\beta N_\gamma} + \mathrm{d} \cdot \mathrm{CO_2} + \mathrm{e} \cdot \mathrm{H_2O} + \mathrm{f} \cdot \mathrm{C_jH_kO_lN_m}$$

Equation 10-4

Where:

- $\mathrm{C_wH_xO_yN_z}$ represents the molecular formula of the carbon source used
- $\mathrm{CH_\alpha O_\beta N_\gamma}$ represents the molecular formula of the biomass produced[19]
- $\mathrm{C_jH_kO_lN_m}$ represents the molecular formula of the product expressed[20]

[19] In the case of Pichia pastoris this is found to be $CH_{1.761}N_{0.143}O_{0.636}S_{0.002}$ at atmospheric oxygen levels (Carnicer et al., 2009)

[20] Based on amino acid sequence and assuming negligible levels of glycosylation, for tHBc-k1,k1 this was determined to be $CH_{1.902}N_{0.284}O_{0.302}S_{0.007}$

10.4.2 Design space for growth kinetics screening in microtitre plates

Table 10-2 Growth kinetics screening in microtitre plates: Design space (left) and corresponding response values (right).

Pattern	V_w(mL)	$[X]_0$ (OD_{600})	t (h)	*$[X]_t$ (OD_{600})*	*μ (h^{-1})*
−−−	0.5	20	2	*27.2*	$1.54 \cdot 10^{-1}$
−−+	0.5	20	30	*58.4*	$3.57 \cdot 10^{-2}$
−+−	0.5	100	2	*140*	$1.70 \cdot 10^{-1}$
−++	0.5	100	30	*108*	$2.44 \cdot 10^{-3}$
+−−	2.5	20	2	*21.2*	$2.91 \cdot 10^{-1}$
+−+	2.5	20	30	*62.8*	$3.81 \cdot 10^{-1}$
++−	2.5	100	2	*105*	$2.53 \cdot 10^{-2}$
+++	2.5	100	30	*98.4*	$-5.38 \cdot 10^{-4}$
000	1.5	60	16	*94.8*	$2.86 \cdot 10^{-2}$
000	1.5	60	16	*95.6*	$2.91 \cdot 10^{-2}$
000	1.5	60	16	*94.8*	$2.86 \cdot 10^{-2}$

10.4.3 Calculation for correlation between incubation time and working volume

$$\mu = 4.91 \cdot 10^{-2} - 3.78 \cdot 10^{-2} f(t) - 3.37 \cdot 10^{-2} f(V_W) + 3.35 \cdot 10^{-2} f(V_W \cdot t) = 0$$

$$\therefore 1.46 - 1.13 f(t) - 1.01 f(V_W) + f(V_W) \cdot f(t) = 0$$

$$\therefore (f(t) - 1.01)(f(V_W) - 1.13) = -0.319$$

$$\therefore f(t) = \frac{-0.319}{f(V_W) - 1.13} + 1.01$$

$$f(t) = \frac{t - 16}{14} \cup f(V_W) = V_W - 1.5$$

$$\therefore \mathbf{t_{(\mu=0)}} = \frac{\mathbf{-4.47}}{\mathbf{V_W - 2.63}} + \mathbf{30.14} \qquad \textbf{Equation 10-5}$$

10.4.4 Design space and model summaries for induction screenings in microtitre plates

Table 10-3 Growth and expression kinetics screening in microtitre plates during methanol induction of *P. pastoris* Mut$^+$: Design space (left) and corresponding response values (right).

Pattern	*[MeOH] (% vol.)*	*$[X]_0$ (OD_{600})*	*t (h)*	*$[X]_t$ (OD_{600})*	*μ (h^{-1})*	*[tHBc] (ng/mL)*	*$tHBc_s$ ($\mu g/mL/OD_{600}$)*
−−−	0.5	18	2	5.54	$-5.89 \cdot 10^{-1}$	8.12	1.47
+−−	5	18	2	6.50	$-5.09 \cdot 10^{-1}$	24.7	0.380
−+−	0.5	90	2	65.6	$-1.58 \cdot 10^{-1}$	25.5	0.389
++−	5	90	2	63.2	$-1.77 \cdot 10^{-1}$	27.1	0.429
0	2.75	54	16	29.2	$-3.84 \cdot 10^{-2}$	76.2	2.61
0	2.75	54	16	31.6	$-3.35 \cdot 10^{-2}$	76.6	2.42
0	2.75	54	16	27.2	$-4.29 \cdot 10^{-2}$	80.0	2.94
−−+	0.5	18	30	10.0	$-1.96 \cdot 10^{-2}$	89.7	8.96
+−+	5	18	30	12.8	$-1.14 \cdot 10^{-2}$	85.2	6.66
−++	0.5	90	30	41.6	$-2.57 \cdot 10^{-2}$	38.3	0.920
+++	5	90	30	53.2	$-1.75 \cdot 10^{-2}$	61.7	1.16

Table 10-4 Growth and expression kinetics screening in microtitre plates during methanol induction of *P. pastoris* Mut^S: Design space (left) and corresponding response values (right).

Pattern	*[MeOH] (% vol.)*	*$[X]_0$ (OD_{600})*	*t (h)*	*$[X]_t$ (OD_{600})*	*μ (h^{-1})*	*[tHBc] (ng/mL)*	*$tHBc_s$ ($\mu g/mL/OD_{600}$)*
−−−	0.3	18	2	6.26	$-5.28 \cdot 10^{-1}$	0	0
+−−	1.0	18	2	6.48	$-5.11 \cdot 10^{-1}$	0	0
−+−	0.3	90	2	79.2	$-6.39 \cdot 10^{-2}$	0	0
++−	1.0	90	2	87.2	$-1.58 \cdot 10^{-2}$	0	0
0	0.65	54	16	26.8	$-4.38 \cdot 10^{-2}$	0	0
0	0.65	54	16	27.2	$-4.29 \cdot 10^{-2}$	0	0
0	0.65	54	16	31.6	$-3.35 \cdot 10^{-2}$	0	0
−−+	0.3	18	30	9.6	$-2.10 \cdot 10^{-2}$	0	0
+−+	1.0	18	30	12.8	$-1.14 \cdot 10^{-2}$	0	0
−++	0.3	90	30	46.4	$-2.21 \cdot 10^{-2}$	13.7	0.295
+++	1.0	90	30	56	$-1.58 \cdot 10^{-2}$	9.02	0.161

Table 10-5 Second screening of growth and expression kinetics screening in microtitre plates during methanol induction of *P. pastoris* Mut$^+$: Design space (left) and corresponding response values (right).

Pattern	*[MeOH] (% vol.)*	*$[X]_0$ (OD_{600})*	*t (h)*	*$[X]_t$ (OD_{600})*	*μ (h^{-1})*	*[tHBc] (ng/mL)*	*$tHBc_s$ (μg/mL/OD_{600})*
−−−	2.5	9	16	23.2	$5.92•10^{-2}$	0.506	$2.18•10^{-2}$
+−−	7.5	9	16	4.92	$-3.77•10^{-2}$	1.99	$4.04•10^{-1}$
−+−	2.5	45	16	71.2	$2.87•10^{-2}$	0.0292	$4.11•10^{-4}$
++−	7.5	45	16	63.6	$2.16•10^{-2}$	0.172	$2.71•10^{-3}$
000	5	27	31	50	$2.05•10^{-2}$	0.445	$8.90•10^{-3}$
000	5	27	31	53.6	$2.29•10^{-2}$	0.450	$8.40•10^{-3}$
000	5	27	31	49.2	$2.00•10^{-2}$	0.0841	$1.71•10^{-3}$
−−+	2.5	9	46	28.8	$2.53•10^{-2}$	0.0499	$1.73•10^{-3}$
+−+	7.5	9	46	24.4	$2.17•10^{-2}$	0.399	$1.63•10^{-3}$
−++	2.5	45	46	72	$1.02•10^{-2}$	0.387	$5.38•10^{-3}$
+++	7.5	45	46	74.4	$1.09•10^{-2}$	0.215	$2.89•10^{-3}$

Table 10-6 Second screening of growth and expression kinetics in microtitre plates during methanol induction of *P. pastoris* Mut^S: Design space (left) and corresponding response values (right).

Pattern	*[MeOH] (% vol.)*	*$[X]_0$ (OD_{600})*	*t (h)*	*$[X]_t$ (OD_{600})*	*μ (h^{-1})*	*[tHBc] (ng/mL)*	*$tHBc_s$ (μg/mL/OD_{600})*
−−−	0.5	90	16	96	$4.03•10^{-3}$	11.4	$1.19•10^{-1}$
+−−	2.5	90	16	109.2	$1.21•10^{-2}$	7.13	$6.53•10^{-2}$
−+−	0.5	135	16	116.8	$-9.05•10^{-3}$	4.55	$3.90•10^{-2}$
++−	2.5	135	16	134	$-4.65•10^{-4}$	6.85	$5.11•10^{-2}$
000	1.5	112.5	31	114.8	$6.75•10^{-4}$	8.10	$7.05•10^{-2}$
000	1.5	112.5	31	114.8	$6.75•10^{-4}$	10.7	$9.35•10^{-2}$
000	1.5	112.5	31	109.2	$-9.92•10^{-4}$	9.05	$8.29•10^{-2}$
−−+	0.5	90	46	93.2	$7.60•10^{-4}$	10.8	$1.16•10^{-1}$
+−+	2.5	90	46	95.2	$1.22•10^{-3}$	8.49	$8.91•10^{-2}$
−++	0.5	135	46	118.8	$-2.78•10^{-3}$	5.95	$5.01•10^{-2}$
+++	2.5	135	46	137.2	$3.51•10^{-4}$	8.31	$6.06•10^{-2}$

Table 10-7 Growth and expression kinetics screening in microtitre plates during PDM induction of *P. pastoris* Mut$^+$: Design space (left) and corresponding response values (right).

Pattern	$[X]_0$ (OD_{600})	*t (h)*	$[X]_t$ (OD_{600})	μ (h^{-1})	*[tHBc]* (μg/mL)	$tHBc_s$ (μg/mL/OD_{600})
−−	10.8	16	18	$3.19 \cdot 10^{-2}$	0.975	0.0542
+−	54	16	58	$4.47 \cdot 10^{-3}$	0	0
00	32.4	30	36	$3.51 \cdot 10^{-3}$	0	0
00	32.4	30	35.6	$3.14 \cdot 10^{-3}$	0	0
−+	10.8	46	16.4	$9.08 \cdot 10^{-3}$	2.34	0.142
++	54	46	56.8	$1.10 \cdot 10^{-3}$	0	0

Table 10-8 Growth and expression kinetics screening in microtitre plates during PDM induction of *P. pastoris* Mut^S: Design space (left) and corresponding response values (right).

Pattern	$[X]_0$ (OD_{600})	*t (h)*	$[X]_t$ (OD_{600})	μ (h^{-1})	*[tHBc]* (μg/mL)	$tHBc_s$ (μg/mL/OD_{600})
−−	108	16	100.8	$-4.31 \cdot 10^{-3}$	3.13	0.0311
+−	162	16	146.4	$-6.33 \cdot 10^{-3}$	4.85	0.0331
00	135	30	120	$-3.93 \cdot 10^{-3}$	9.60	0.0800
00	135	30	118	$-4.49 \cdot 10^{-3}$	10.8	0.0911
−+	108	46	98	$-2.11 \cdot 10^{-3}$	10.6	0.108
++	162	46	147.6	$-2.02 \cdot 10^{-3}$	6.16	0.0417

Table 10-9: Relative factor contributions in a two-factor, two-level, full-fractional factorial screening designs investigating induction of *P. pastoris* Mut$^+$ in microtitre plates using defined media containing pectin derivatives. *: Curvature detected in screening response model.

Factor		$[X]_0$ (OD_{600})	*t (h)*
Range		*(10.8,54.0)*	*(16, 46)*
% contribution	μ (h^{-1})	60.53%	39.47%
	[tHBc](μg/mL)	77.59%	22.41%
	$tHBc_s$(μg/mL/OD_{600})	74.89%	25.11%

Table 10-10 Relative factor contributions in a two-factor, two-level, full-fractional factorial screening designs investigating induction of *P. pastoris* Mut^S in microtitre plates using defined media containing pectin derivatives. *: Curvature detected in screening response model.

Factor		$[X]_0$ (OD_{600})	*t (h)*
Range		*(108,162)*	*(16, 46)*
% contribution	μ (h^{-1})	14.72%	85.28%
	[tHBc](μg/mL)	29.32%	70.68%
	$tHBc_s$(μg/mL/OD_{600})	43.32%	56.68%

Table 10-11 dGAME Models and goodness of fit for second screening designs investigating induction of *P. pastoris* Mut$^+$ in microtitre plates using defined media containing pectin derivatives.

	All terms (Complex model)			Significant terms (Simplified model)	
	R^2	***R^2 adj.***	***Formulae***	***R^2***	***R^2 adj.***
μ (h^{-1})	0.863	0.657	$\mu = 8.87 \cdot 10^{-3} - 8.70 \cdot 10^{-3} f([X_0])$	0.467	0.333
[tHBc] (μg/mL)	0.801	0.501	$[tHBc] = 0.552 - 0.828 f([X_0])$	0.598	0.498
[tHBc] (μg/mL/OD_{600})	0.808	0.520	$HBc_S = 3.27 \cdot 10^{-2} - 4.92 \cdot 10^{-2} f([X_0])$	0.576	0.470

$$f([X_0]) = \frac{[X_0] - 32.4}{21.6}$$

Table 10-12 dGAME Models and goodness of fit for second screening designs investigating induction of *P. pastoris* Mut^S in microtitre plates using defined media containing pectin derivatives.

	All terms (Complex model)			Significant terms (Simplified model)	
	R^2	***R^2 adj.***	***Formulae***	***R^2***	***R^2 adj.***
μ (h^{-1})	0.972	0.929	$\mu = -3.83 \cdot 10^{-3} - 1.64 \cdot 10^{-3} f([t])$	0.816	0.770
[tHBc] (μg/mL)	0.551	-0.124	$[tHBc] = 7.56 - 2.10 f([t])$	0.337	0.171
[tHBc] (μg/mL/OD_{600})	0.715	0.285	$HBc_S = 6.46 \cdot 10^{-2} - 2.06 \cdot 10^{-2} f([t])$	0.312	0.141

$$f([X_0]) = \frac{[X_0] - 135}{27} \qquad f(t) = \frac{t - 31}{15}$$

10.4.5 Reaction mechanism of d-galacturonic acid and methanol

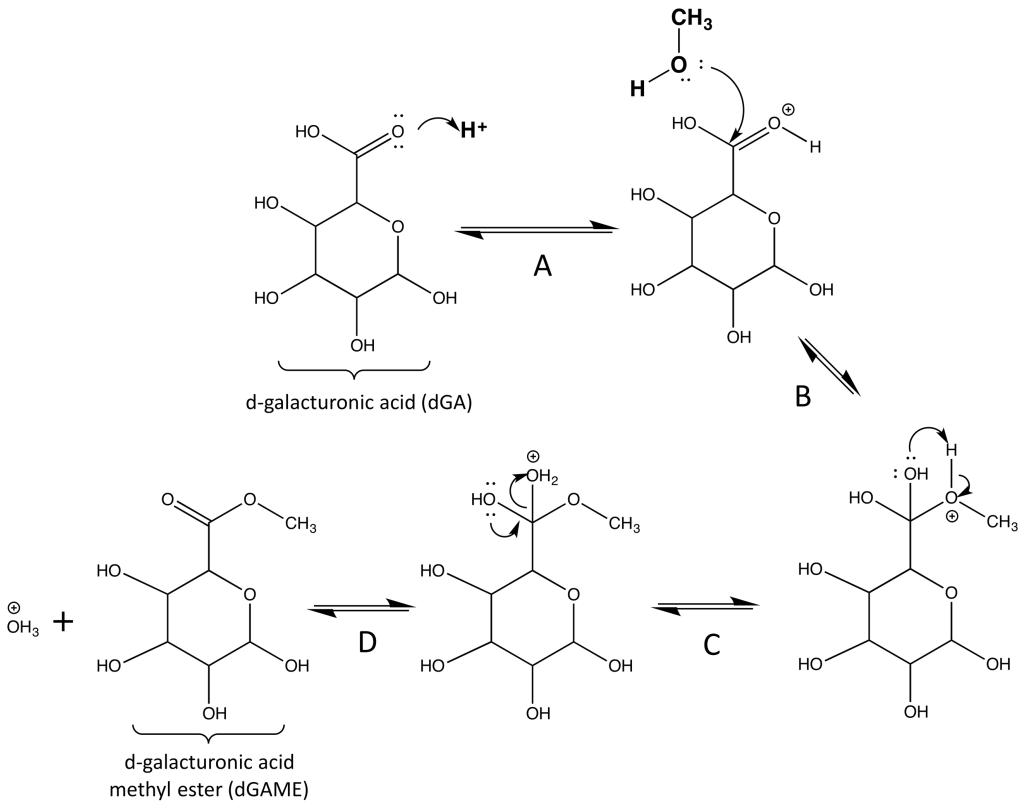

Figure 10.10 Reaction mechanism of d-galacturonic acid and methanol under acidic conditions

10.5 Supplementary data (Chapter 6)

10.5.1 Volumetric total core protein

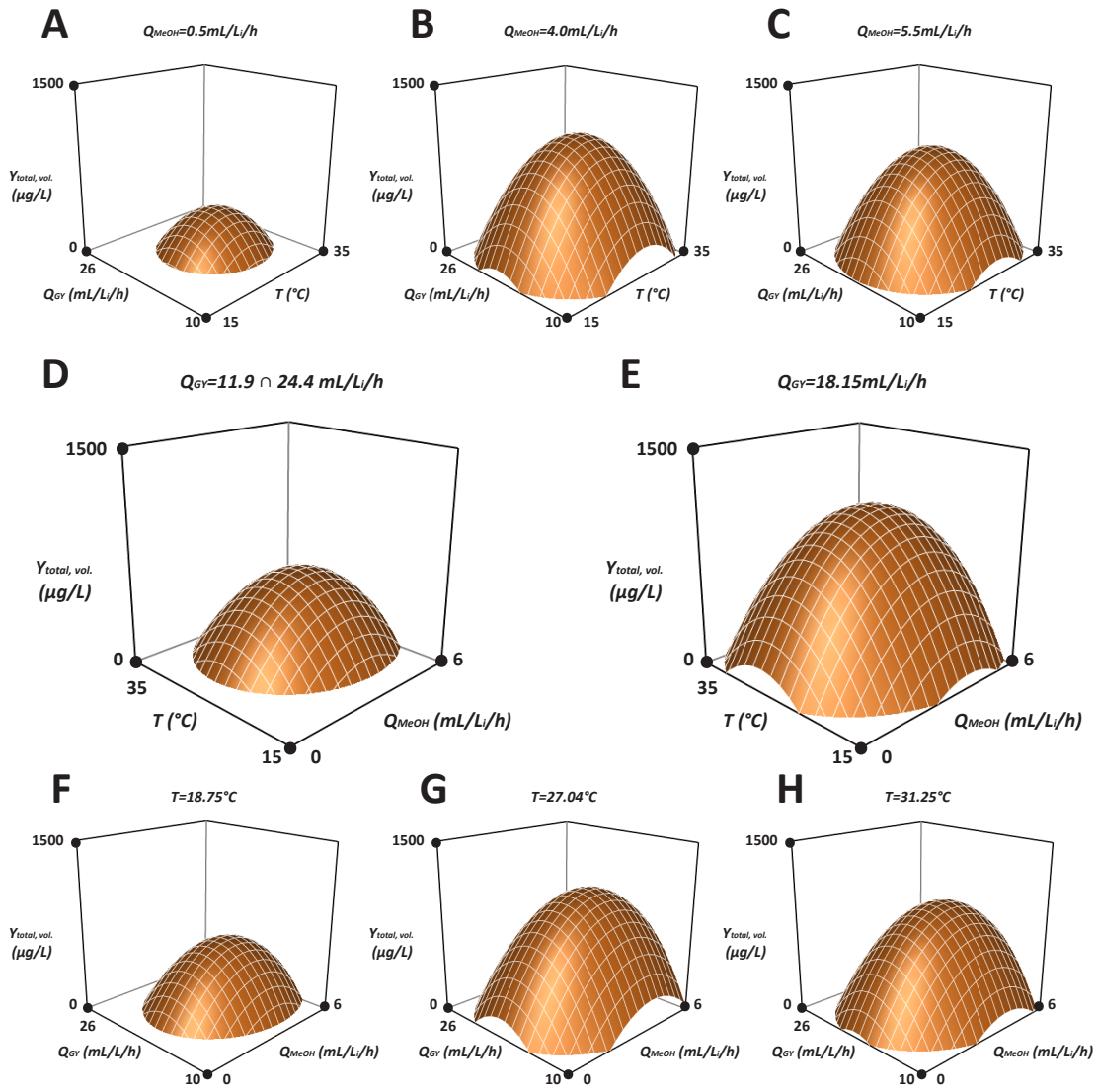

Figure 10.11 Response Surface Model for volumetric total core protein expression, $Y_{total, vol}$. (A-C) Effect of Q_{GY} and T at (A) Q_{MeOH}=0.5mL/L$_i$/h, (B) Q_{MeOH}=3.65mL/L$_i$/h (response maximum) and (C) Q_{MeOH}=5.5mL/L$_i$/h. (D-E) Effect of Q_{MeOH} and T at (D) Q_{GY}=11.90mL/L$_i$/h and Q_{GY}=24.40mL/L$_i$/h; and (E) Q_{GY}=18.15mL/L$_i$/h (response maximum). (F-H) Effect of Q_{GY} and Q_{MeOH} at (F) T=18.75°C, (G) T=27.04°C and (H) T=31.25°C.

10.5.2 Volumetric expression responses

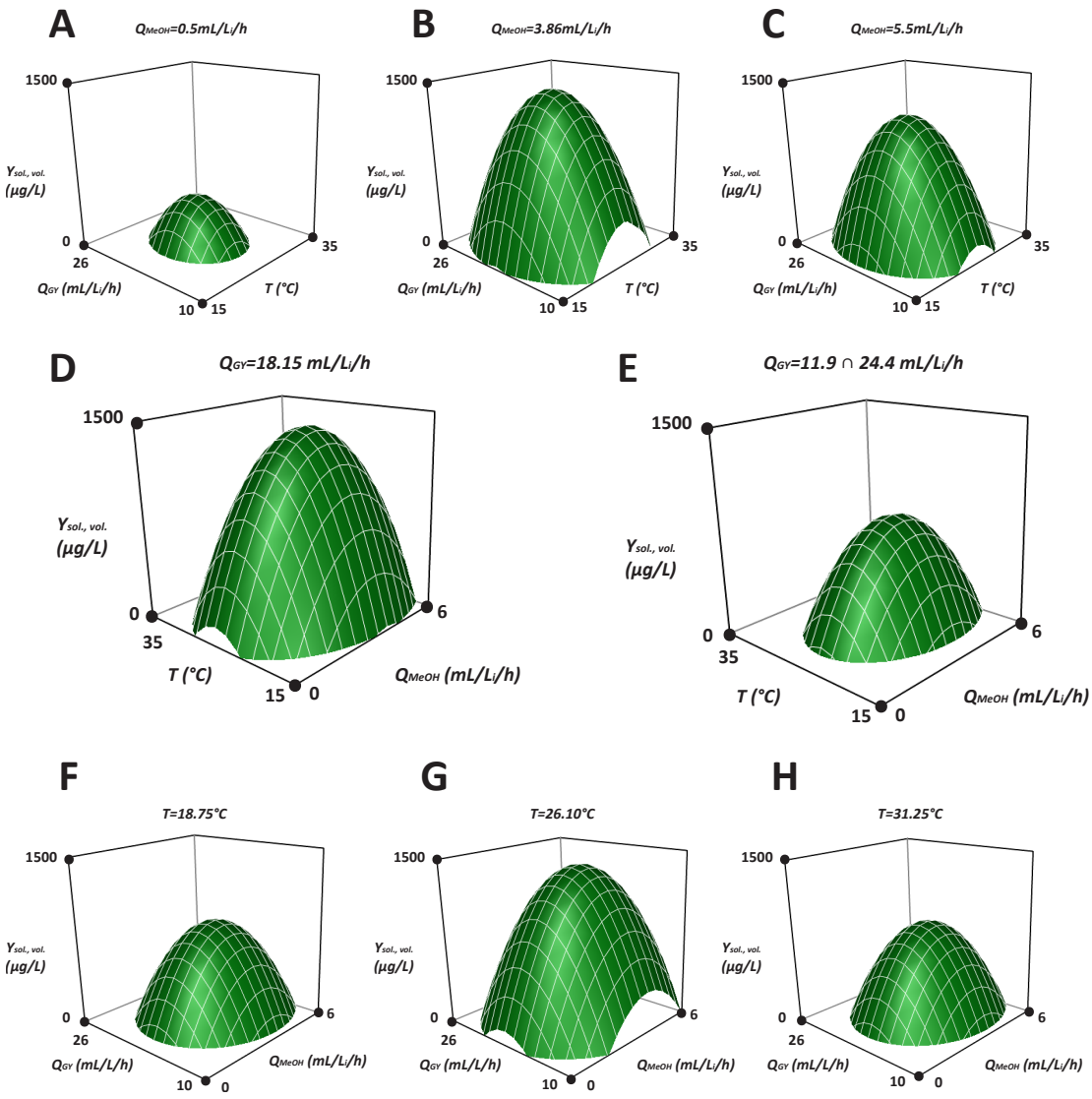

Figure 10.12 Response Surface Model for volumetric soluble core protein expression, $Y_{sol, vol}$. (A-C) Effect of Q_{GY} and T at (A) Q_{MeOH}=0.5mL/Lᵢ/h, (B) Q_{MeOH}=3.86mL/Lᵢ/h (response maximum) and (C) Q_{MeOH}=5.5mL/Lᵢ/h. (D-E) Effect of Q_{MeOH} and T at (D) Q_{GY}=18.15mL/Lᵢ/h (response maximum); and (E) Q_{GY}=11.90mL/Lᵢ/h and Q_{GY}=24.40mL/Lᵢ/h. (F-H) Effect of Q_{GY} and Q_{MeOH} at (F) T=18.75°C, (G) T=26.10°C and (H) T=31.25°C.

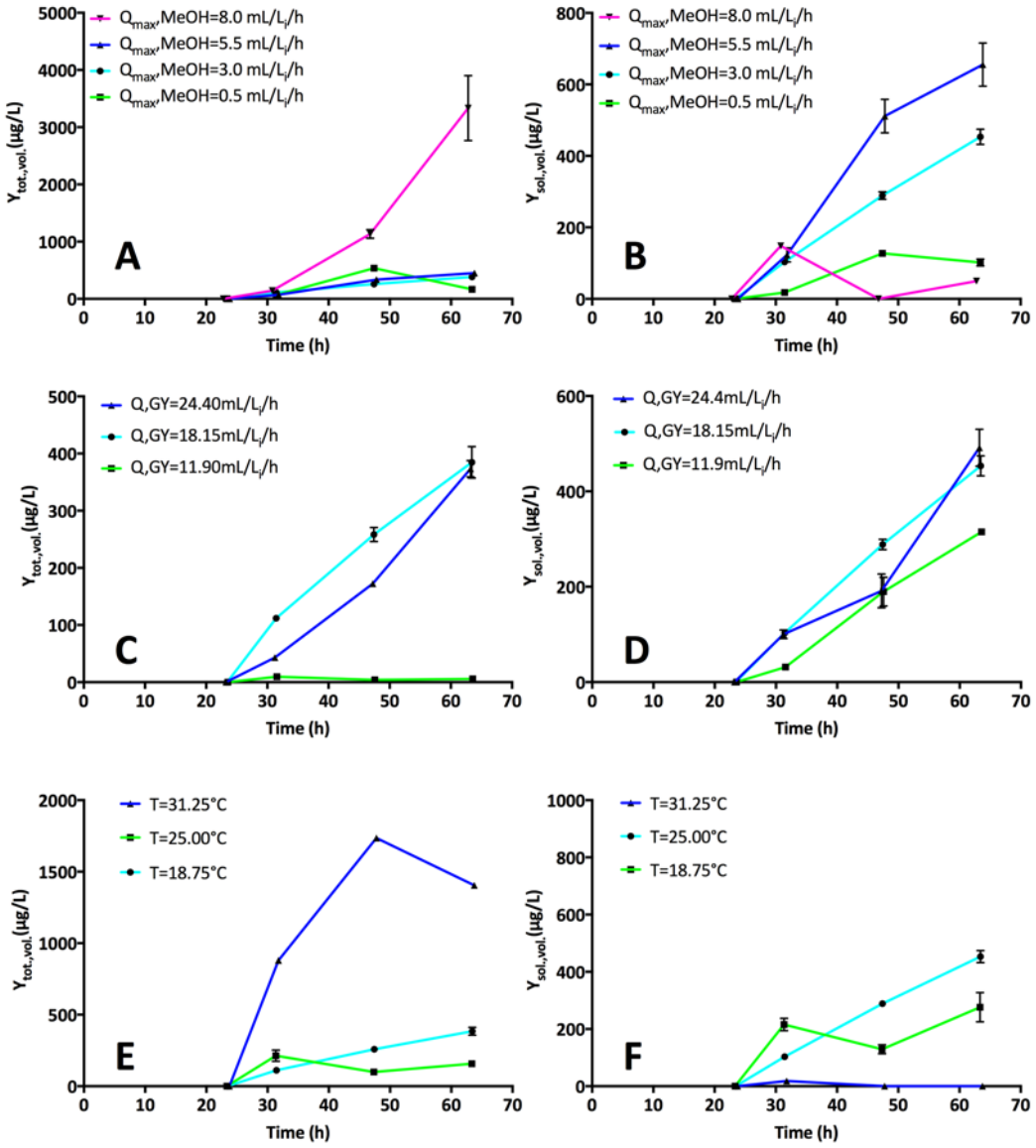

Figure 10.13 Time course profiles of volumetric tHBc expression: (A and B) Expression at various methanol flow rates of total and soluble tHBc respectively. (C and D) Expression at various glycerol flow rates of total and soluble tHBc respectively. (E and F) Expression at various induction temperatures of total and soluble tHBc respectively.

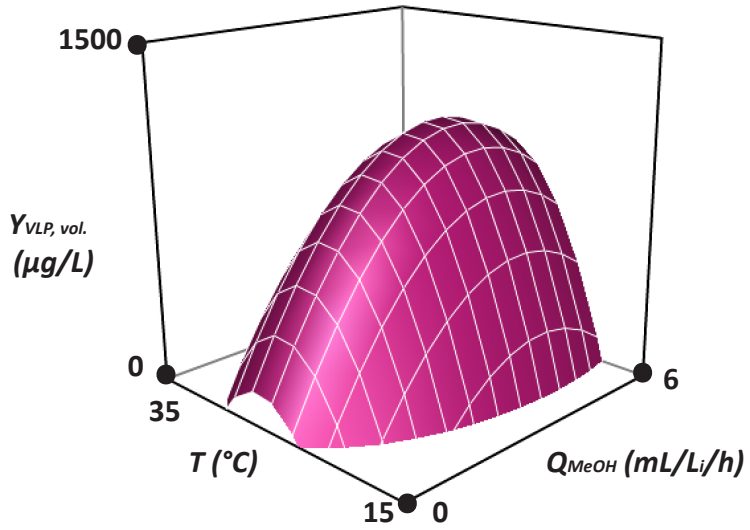

Figure 10.14 Response Surface Model for volumetric VLP core protein expression, $Y_{VLP, vol}$. Effect of Q_{MeOH} and T. Effect of Q_{GY} was found to be negligible.

Table 10-13: Central composite design for induction optimisation. Design space (left) and corresponding response values (right).

Design space					Metabolic responses			*Specific expression (µg/gWCW)*			*Volumetric expression(µg/L)*		
Standard order	**Pattern**	**Q_{MeOH} (mL/L_i/h)**	**Q_{GY} (mL/L_i/h)**	**T (°C)**	**[X] (g_{WCW}/L)**	**CER (mmol/L/h)**	**RQ**	$Y_{tot.,\ spec.}$	$Y_{sol.,\ spec.}$	$Y_{VLP,\ spec.}$	$Y_{tot.,\ vol.}$	$Y_{sol.,\ vol.}$	$Y_{VLP.,\ vol.}$
1	−−−	1	13.15	20	180	16.7	0.568	$1.93\cdot10^{-1}$	$1.87\cdot10^{-2}$	$3.76\cdot10^{-1}$	34.7	3.36	67.6
2	−−+	1	13.15	30	181	18.5	0.570	1.43	$4.05\cdot10^{-1}$	2.58	259	73.2	466
3	−+−	1	23.15	20	203	14.6	0.647	$6.70\cdot10^{-2}$	$3.05\cdot10^{-2}$	$1.67\cdot10^{-1}$	13.6	6.19	33.8
4	−++	1	23.15	30	209	18.6	0.675	$5.39\cdot10^{-1}$	$5.39\cdot10^{-1}$	1.32	113	113	275
5	+−−	5	13.15	20	257	78.5	0.522	$9.09\cdot10^{-1}$	$9.82\cdot10^{-2}$	2.09	234	25.3	537
6	+−+	5	13.15	30	139	28.0	0.589	2.60	2.42	3.13	361	336	435
7	++−	5	23.15	20	270	69.4	0.611	1.33	1.28	$8.08\cdot10^{-1}$	360	347	218
8	+++	5	23.15	30	162	40.7	0.704	$7.85\cdot10^{-1}$	1.18	1.18	127	192	191
9	000	3	18.15	25	238	45.7	0.551	3.56	3.79	2.93	845	900	696
10	000	3	18.15	25	240	46.5	0.574	2.75	3.95	3.17	660	949	762
11	000	3	18.15	25	228	45.3	0.546	3.14	4.96	4.12	714	$1.13\cdot10^{3}$	937
12	0α0	3	24.4	25	250	48.9	0.558	2.57	4.08	4.34	642	$1.02\cdot10^{3}$	$1.08\cdot10^{3}$
13	α00	5.5	18.15	25	279	76.7	0.534	5.85	7.87	7.52	$1.63\cdot10^{3}$	$2.19\cdot10^{3}$	$2.10\cdot10^{3}$
14	0-α0	3	11.9	25	217	45.8	0.546	4.00	5.58	5.45	870	$1.21\cdot10^{3}$	$1.18\cdot10^{3}$
15	00α	3	18.15	31.25	178	31.7	0.632	8.37	9.65	7.49	$1.49\cdot10^{3}$	$1.72\cdot10^{3}$	$1.33\cdot10^{3}$
16	00-α	3	18.15	18.75	228	44.9	0.538	1.97	2.77	1.67	448	63.1	381
17	-α00	0.5	18.15	25	168	9.86	0.655	7.97	1.06	$6.02\cdot10^{-1}$	134	178	101

10.6 Supplementary data (Chapter 7)

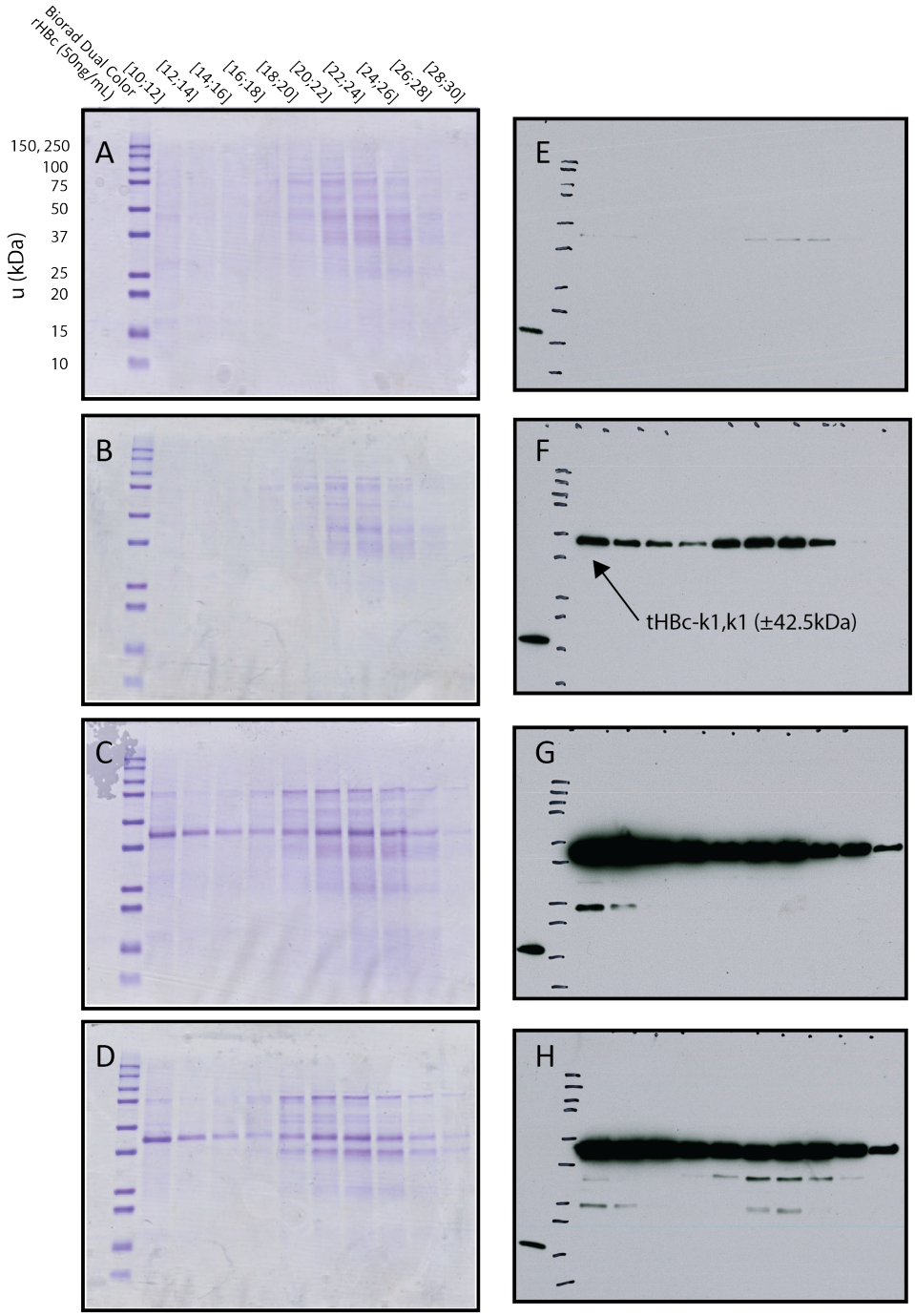

Figure 10.15 Size Exclusion Chromatography (2.4.11) fractions of bead-lysed (2.4.9) k1, k1 construct (42.5kDa) using pure methanol induction. Two-minute chromatography fractions were collected after 10 minutes for an additional 20 minutes. (A-D) Coomassie stains of reduced gels. (E-H) Western Blots (1-minute exposure). Fermentation samples were collected at 0 hours (A & E), 8 hours (B & F), 24 hours (C & G) and 40 hours (D & H) of induction. SEC and Western blotting procedures were performed by iQur Ltd.

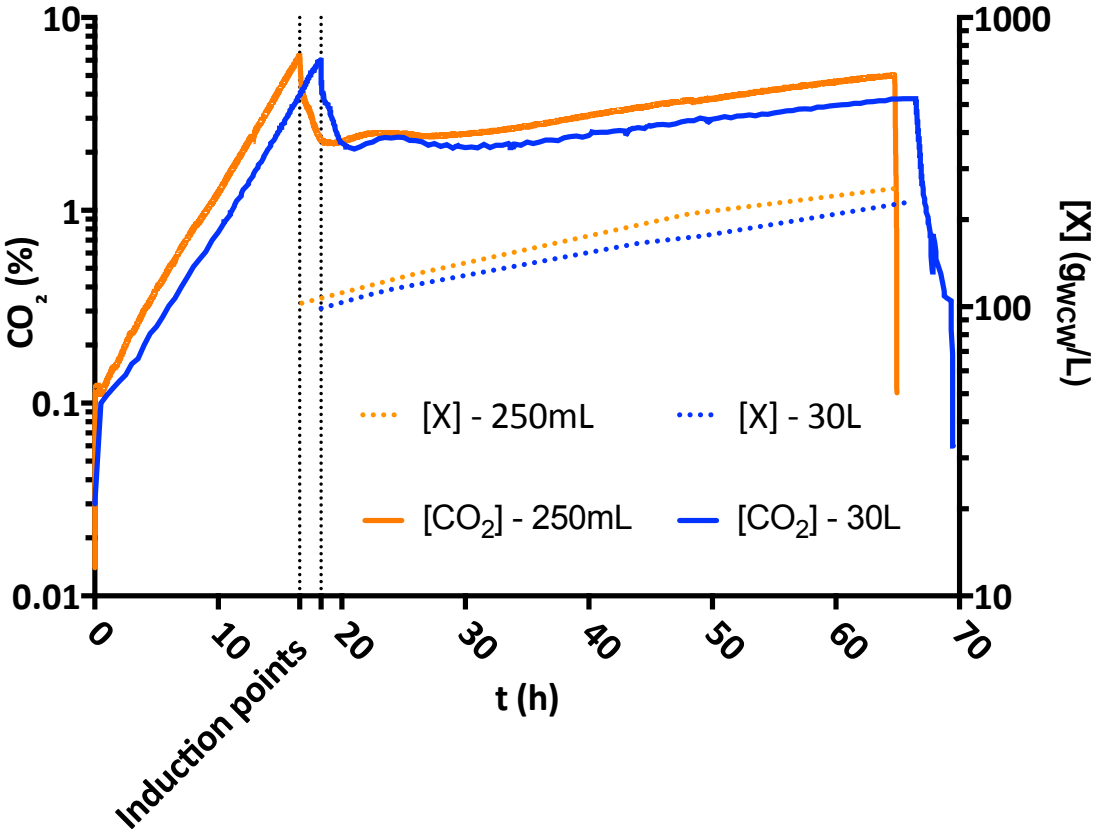

Figure 10.16 Log-scaled fermentation data of process scale-up of *Pichia pastoris* expressing tHBc-HA2.3, $(M2e)_3$ using mixed feeding induction. Large-scale data, at 30L total working volume, is shown in blue. Small-scale data, at 250mL total working volume, is shown in orange. Carbon dioxide concentrations ([CO_2]) was measured in off-gas and biomass concentrations ([X]) were determined through wet cell weight analysis.